

# SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: 02-02-04  
Searcher: Beverly C #2528  
Terminal time: 20  
Elapsed time: \_\_\_\_\_  
CPU time: \_\_\_\_\_  
Total time: 23  
Number of Searches: \_\_\_\_\_  
Number of Databases: 1

### Search Site

\_\_\_\_\_ STIC  
\_\_\_\_\_ CM-1  
\_\_\_\_\_ Pre-S

### Type of Search

\_\_\_\_\_ N.A. Sequence  
\_\_\_\_\_ A.A. Sequence  
\_\_\_\_\_ Structure  
\_\_\_\_\_ Bibliographic

### Vendors

\_\_\_\_\_ IG  
\_\_\_\_\_ STN  
\_\_\_\_\_ Dialog  
\_\_\_\_\_ APS  
\_\_\_\_\_ Geninfo  
\_\_\_\_\_ SDC  
\_\_\_\_\_ DARC/Questel  
☒ Other CGN

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2004, 00:03:08 ; Search time 1498 Seconds  
(without alignments)  
15874.988 Million cell updates/sec

Title: US-10-005-338B-1

Perfect score: 6525  
Sequence: 1 aaatgttgatatttctct.....ttgatcataataagtgaat 6525

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6524	100.0	6525	15	US-10-005-338B-1 Sequence 1, Appli
2	5373.2	82.3	5475	14	US-10-090-458-4 Sequence 4, Appli
3	5255.8	80.5	5463	14	US-10-090-458-1 Sequence 1, Appli
4	5247.4	80.4	5262	10	US-09-971-121-5 Sequence 5, Appli
5	4918.4	75.4	4929	10	US-09-971-121-1 Sequence 1, Appli
6	4807.4	73.7	4917	14	US-10-090-458-3 Sequence 3, Appli
7	4754.8	72.9	4785	10	US-09-971-121-3 Sequence 3, Appli
8	3258.2	49.9	3347	12	US-10-108-260A-160 Sequence 160, App
9	3178	48.7	3268	13	US-10-094-749-984 Sequence 984, App
10	1789.4	27.4	3928	11	US-09-822-846-98 Sequence 98, Appl
11	1731.8	26.5	1943	12	US-10-108-260A-1424 Sequence 1424, Ap
12	1722.8	26.4	1818	11	US-10-204-887-32 Sequence 32, Appl
13	1543.8	23.7	1548	15	US-09-822-846-97 Sequence 97, Appl
14	1182	18.1	1346	9	US-09-962-436-337 Sequence 337, App
15	1182	18.1	1346	10	US-09-880-107-3406 Sequence 3406, Ap

16	1182	18.1	1346	10	US-09-954-531-1033	Sequence 1033, Ap	
17	1182	18.1	1346	15	US-10-171-581-333	Sequence 333, App	
18	1150.4	17.6	6181	15	US-10-005-338B-4	Sequence 4, Appli	
19	1118.8	17.1	5981	15	US-10-005-338B-3	Sequence 3, Appli	
20	1117.2	17.1	4875	14	US-10-090-454-3	Sequence 3, Appli	
21	1117.2	17.1	5018	14	US-10-090-454-1	Sequence 1, Appli	
22	1090.8	16.7	5296	15	US-10-005-338B-2	Sequence 2, Appli	
23	1086	16.6	4854	14	US-10-090-453A-3	Sequence 3, Appli	
24	1086	16.6	5332	14	US-10-090-453A-1	Sequence 1, Appli	
25	908.2	13.9	4766	13	US-10-093-463-199	Sequence 199, App	
26	894.8	13.7	4798	13	US-10-093-463-197	Sequence 197, App	
27	797.4	12.2	5175	12	US-10-085-198-37	Sequence 37, Appl	
28	757.4	11.6	974	10	US-09-833-381-1085	Sequence 1085, Ap	
29	670.4	10.3	740	12	US-10-242-355-968	Sequence 968, App	
30	670.4	10.3	740	12	US-10-242-355-969	Sequence 969, App	
31	651.4	10.0	2645	13	US-10-094-749-1290	Sequence 1290, Ap	
32	600	9.2	668	12	US-10-242-355-970	Sequence 970, App	
33	481.8	7.4	2512	12	US-10-108-260A-1795	Sequence 1795, Ap	
C	34	475.4	7.3	477	10	US-09-867-701-10804	Sequence 10804, A
35	471.2	7.2	476	10	US-09-867-701-4362	Sequence 4362, Ap	
36	459	7.0	465	12	US-10-242-535A-44227	Sequence 44227, A	
37	452.8	6.9	456	12	US-10-242-535A-16105	Sequence 16105, A	
38	410	6.3	473	12	US-10-242-535A-28509	Sequence 28509, A	
C	39	398.6	6.1	412	10	US-09-867-701-4371	Sequence 4371, Ap
40	390.6	6.0	431	12	US-10-264-237-22	Sequence 22, Appl	
41	373.6	5.7	1986	9	US-09-767-870-4	Sequence 4, Appli	
42	373.6	5.7	1986	13	US-10-242-568-4	Sequence 4, Appli	
C	43	369.6	5.7	469	10	US-09-880-107-1883	Sequence 1883, Ap
44	319.8	4.9	2007	12	US-10-264-049-708	Sequence 708, App	
45	319.2	4.9	1350	10	US-09-833-381-1082	Sequence 1082, Ap	

ALIGNMENTS

RESULT 1

US-10-005-338B-1  
; Sequence 1, Application US/10005338B  
; Publication No. US20030044895A1  
; GENERAL INFORMATION:  
; APPLICANT: DENEFELE, Patrice  
; APPLICANT: ROSIER-MONTUS, Marie-Francoise  
; APPLICANT: PRADES, Cathérine  
; APPLICANT: ARNOULD-REGUIGNE, Isabelle  
; APPLICANT: DUVERGER, Nicolas  
; APPLICANT: ALLIEMETS, Rando  
; APPLICANT: DEAN, Michael  
; TITLE OF INVENTION: NUCLEIC ACIDS OF THE HUMAN ABCA5, ABCA6, ABCA9, AND ABCA10 GENES,  
; FILE REFERENCE: ABCA5, 6, 9, 10  
; CURRENT APPLICATION NUMBER: US/10/005,338B  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: US 60/263,231  
; PRIOR FILING DATE: 2001-01-23  
; PRIOR APPLICATION NUMBER: FR 00403440.1  
; PRIOR FILING DATE: 2000-12-07  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 6525  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 4449  
; OTHER INFORMATION: n=unknown, may be a o r c o r t  
US-10-005-338B-1

Query Match 100.0%; Score 6524; DB 15; Length 6525;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 6525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAAATGTTGATATTTCTCTTAGCAGGCTGTCAACCAGGTTAGTTTCAGGTCAATAAGTTT 60

Db	1	AAAAATGTTGATATATTTCTCTTAGCAGCGTGTCAACCAAGGTTAGGTTTCAGGTCTATAAGTTT	60
Qy	61	CTACCCACATTTCTTTGAACGTAGTGTGTCATTTTAGTATTATTTTCAAAACATTTTGCAG	120
Db	61	CTACCCACATTTCTTTGAACGTAGTGTGTCATTTTAGTATTATTTTCAAAACATTTTGCAG	120
Qy	121	TACCTTTTGGTCTGTCTTGTTGTGTGCTTGCAGTGAACAGTCTGGATTTGGACAGTGGT	180
Db	121	TACCTTTTGGTCTGTCTTGTTGTGTGCTTGCAGTGAACAGTCTGGATTTGGACAGTGGT	180
Qy	181	CTGTCTGTTAGTTCAGTTTCTCAAGCCTTTGTGTCACATAATAGGATTTGGAATTATGATG	240
Db	181	CTGTCTGTTAGTTCAGTTTCTCAAGCCTTTGTGTCACATAATAGGATTTGGAATTATGATG	240
Qy	241	TCCAGCTTTGGAAATATATACAGGAATATAAAACAACCTTTTAGAGTGCCTTTCCTGAGCTC	300
Db	241	TCCAGCTTTGGAAATATATACAGGAATATAAAACAACCTTTTAGAGTGCCTTTCCTGAGCTC	300
Qy	301	TCTTTCTATTGTTCCCCCTCTACTTTTGTCTTCCCTGTGGCTGTGTTTCTATCCTCC	360
Db	301	TCTTTCTATTGTTCCCCCTCTACTTTTGTCTTCCCTGTGGCTGTGTTTCTATCCTCC	360
Qy	361	AGCCAGAGAGCTAGTGTTATTTTCTCCATTTGTGTTACACTTTGTGCAGCTGCAACAC	420
Db	361	AGCCAGAGAGCTAGTGTTATTTTCTCCATTTGTGTTACACTTTGTGCAGCTGCAACAC	420
Qy	421	CATATCCAGGCCCAATGTTAGGAGGTAGAGAGAAAAGCAAAAGGATTTGGCCTCATCC	480
Db	421	CATATCCAGGCCCAATGTTAGGAGGTAGAGAGAAAAGCAAAAGGATTTGGCCTCATCC	480
Qy	481	TCTTACAACAGATAGTTCATTTGAATAGAGAGAAAGGTTTTCCTGCCCTCAGAGTGTGGCT	540
Db	481	TCTTACAACAGATAGTTCATTTGAATAGAGAGAAAGGTTTTCCTGCCCTCAGAGTGTGGCT	540
Qy	541	GCACCTAGGCTTTTGTGTTACTGTAGTCTGCGCCCTGTTTACCATGGGATTCCTTCCATGTGGG	600
Db	541	GCACCTAGGCTTTTGTGTTACTGTAGTCTGCGCCCTGTTTACCATGGGATTCCTTCCATGTGGG	600
Qy	601	ATACAGAGGAATTCAGAGAAAAGAAAGATTTCCTATTTCTACATTTCTCCCTGAGCATTT	660
Db	601	ATACAGAGGAATTCAGAGAAAAGAAAGATTTCCTATTTCTACATTTCTCCCTGAGCATTT	660
Qy	661	AAGACCTCCCTTGCCCATTCCTCAATTCAAAGCTTAAGCTTCTTCTCGAGCTGCCTCTGT	720
Db	661	AAGACCTCCCTTGCCCATTCCTCAATTCAAAGCTTAAGCTTCTTCTCGAGCTGCCTCTGT	720
Qy	721	GGCGGTTTCGGGAGATACCAAGAGAGAAAAGTACCACTGTTGATATGGTGGTATTTCAA	780
Db	721	GGCGGTTTCGGGAGATACCAAGAGAGAAAAGTACCACTGTTGATATGGTGGTATTTCAA	780
Qy	781	ATTCTGTGCTACCCATTTTTCATGCTTGTGTTACTTTTCAGAGCTGACAGATTCGTGCT	840
Db	781	ATTCTGTGCTACCCATTTTTCATGCTTGTGTTACTTTTCAGAGCTGACAGATTCGTGCT	840
Qy	841	CCATGCAATCTGTCAGTTTCCCTTAAGAGAGACGCTTGGAGTATGCTTAATCCATCTTAC	900
Db	841	CCATGCAATCTGTCAGTTTCCCTTAAGAGAGACGCTTGGAGTATGCTTAATCCATCTTAC	900
Qy	901	CTGGGACTGAAACAGCTGCTTATTTTGCCTGTAAAAATTTACATGCAGTTTACTGCGTGGC	960
Db	901	CTGGGACTGAAACAGCTGCTTATTTTGCCTGTAAAAATTTACATGCAGTTTACTGCGTGGC	960
Qy	961	TCCGGGTTTGTGTTGTTTTTCTCTTTAATAGGTTTATTCAGAAAAAATGTCACATG	1020
Db	961	TCCGGGTTTGTGTTGTTTTTCTCTTTAATAGGTTTATTCAGAAAAAATGTCACATG	1020
Qy	1021	CAATTAGGGAGGTAGGATTTGGAGACAGACCAGAACACTTCTACTGAAGATTTACTTAA	1080
Db	1021	CAATTAGGGAGGTAGGATTTGGAGACAGACCAGAACACTTCTACTGAAGATTTACTTAA	1080
Qy	1081	TTAAATGCAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTTTCCACTATTTTTTTTAT	1140

1081	TTAAATGCGAAGCAAAAAGAGTAGTGTTCACGAAATTCCTTTTCCCATATTTTTTTTAT	1114
1141	TTTGGTTAAATATTAATTTAGCATGATGCATCAAATAAGAAATATGAGAAGTGCCTAATA	1200
1141	TTTGGTTAAATATTAATTTAGCATGATGCATCAAATAAGAAATATGAGAAGTGCCTAATA	1200
1201	TAGAACTCAATCCTATATGGAACAAGTTTACTCTTTCTAAATCTAAATCTTGGATATATCCAG	1260
1201	TAGAACTCAATCCTATATGGAACAAGTTTACTCTTTCTAAATCTAAATCTTGGATATATCCAG	1260
1261	TGACTAAATATTACAAGCAGCATCATCGAAGAGTGTCTACTGATCATCTCAATCTCATGTC	1320
1261	TGACTAAATATTACAAGCAGCATCATCGAAGAGTGTCTACTGATCATCTCAATCTCATGTC	1320
1321	TAAATCTGAAAGATATACAAATCAAAAAGAAATGTTAAACATCCAGTCTCTCTAAGCCGA	1380
1321	TAAATCTGAAAGATATACAAATGAAAAGAAATGTTAAACATCCAGTCTCTCTAAGCCGA	1380
1381	GCAACTTTGTAGGTGTGGTTTTTCAAGACTCCAGTCTCAAGAGCTGGCTGTTCAAAATCANGTG	1440
1381	GCAACTTTGTAGGTGTGGTTTTTCAAGAGCTCCATGTCTTATGAACTTCGTGTTTTTCTCGT	1440
1441	ATATGATTTCCAGTATCTCTTATTTATGGAATTCAGAGCTGGCTGTTCAAAATCATGTG	1500
1441	ATATGATTTCCAGTATCTCTTATTTATGGAATTCAGAGCTGGCTGTTCAAAATCATGTG	1500
1501	AGGCTGCTCAGTACTGTGCTCAGGTTTTACAGTTTTTACAAGCATCCATAGATGCTGCCA	1560
1501	AGGCTGCTCAGTACTGTGCTCAGGTTTTACAGTTTTTACAAGCATCCATAGATGCTGCCA	1560
1561	TTATACAGTTTGAAGCAAAATGTTTTCTCTTTGGAAGAGCTGGAGTCAACTAAAGCTGTTA	1620
1561	TTATACAGTTTGAAGCAAAATGTTTTCTCTTTGGAAGAGCTGGAGTCAACTAAAGCTGTTA	1620
1621	TTATGGGAGAACTGCTGTTGTAGAAAATAGATACCTTTCCCGGAGAGTAATTTTAATAT	1680
1621	TTATGGGAGAACTGCTGTTGTAGAAAATAGATACCTTTCCCGGAGAGTAATTTTAATAT	1680
1681	ACCTAGTTTATAGCATTTTACCTTTTGGATACCTTTTGGCAATTCATATCGTAGCAGAAA	1740
1681	ACCTAGTTTATAGCATTTTACCTTTTGGATACCTTTTGGCAATTCATATCGTAGCAGAAA	1740
1741	AAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACCTGCTTTTGGC	1800
1741	AAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACCTGCTTTTGGC	1800
1801	TTTCTCGGGTCTCTATATACAAGTTTAAATTTTCTTATGTCCTCTTATGCGAGTCA	1860
1801	TTTCTCGGGTCTCTATATACAAGTTTAAATTTTCTTATGTCCTCTCTTATGCGAGTCA	1860
1861	TTGGCAGAGCTTCTTGTTATTTCTCAAGTAGCAGCATCTGATATTTCTGCTTTTTT	1920
1861	TTGGCAGAGCTTCTTGTTATTTCTCAAGTAGCAGCATCTGATATTTCTGCTTTTTT	1920
1921	TGCTTTATGGATATCATCTGTAATTTTTTGTCTTAAATGCTGACACCTCTTTTAAAAAAT	1980
1921	TGCTTTATGGATATCATCTGTAATTTTTTGTCTTAAATGCTGACACCTCTTTTAAAAAAT	1980
1981	CAAAAATGCTGGGAATAGTTGAAATTTTTTGTACTGCTGGCTTTTGGATTTATGCGCTTA	2040
1981	CAAAAATGCTGGGAATAGTTGAAATTTTTTGTACTGCTGGCTTTTGGATTTATGCGCTTA	2040
2041	TGTAATCCTCATAGNAAGTTTTTCCCAAAATCGTTAGTGGGCTTTTCACTCTTCTGTC	2100
2041	TGTAATCCTCATAGNAAGTTTTTCCCAAAATCGTTAGTGGGCTTTTCACTCTTCTGTC	2100
2101	ACTGCTATTTTGTATTTGTTATTTGCAAGCTCATGATTTAGAGATTTTAATGAAGGTG	2160
2101	ACTGCTATTTTGTATTTGTTATTTGCAAGCTCATGCAITTTAGAGATTTTAATGAAGGTG	2160
2161	CTTCAATTTCAAAATTTGACTGCGAGGCCATATCCTCTAAATTAATTAACAATATCATGCTCA	2220
2161	CTTCAATTTCAAAATTTGACTGCGAGGCCATATCCTCTAAATTAATTAACAATATCATGCTCA	2220



2221 CACTTAATAGTATATTTCTATGTCTCTTGGCTGTCTATCTTGTATCAAGTCATTTCCAGGGG 2280  
Db |||||  
2221 CACTTAATAGTATATTTCTATGTCTCTTGGCTGTCTATCTTGTATCAAGTCATTTCCAGGGG 2280  
Qy |||||  
2281 AATTTGGCTTACGGAGATCATCTTTATATTTCTTGAAGCTTCAATATTTGGTCAAGAGTA 2340  
Db |||||  
2281 AATTTGGCTTACGGAGATCATCTTTATATTTCTTGAAGCTTCAATATTTGGTCAAGAGTA 2340  
Qy |||||  
2341 AAAGAAATTAAGGAGTATACAGAGGCAATGTTAATGGAATATTTAGTTAGTAA 2400  
Db |||||  
2341 AAAGAAATTAAGGAGTATACAGAGGCAATGTTAATGGAATATTTAGTTAGTAA 2400  
Qy |||||  
2401 TTATTGAGCCAGTTCTTTCAGAAATTTGTAGAAAGAGCCATAAGAAATTTAGTGGTATTC 2460  
Db |||||  
2401 TTATTGAGCCAGTTCTTTCAGAAATTTGTAGAAAGAGCCATAAGAAATTTAGTGGTATTC 2460  
Qy |||||  
2461 AGAAGACATACAGAAAGAGGGTGAATATGTGAGGCTTTTGAAATTTGTCTATTTGACA 2520  
Db |||||  
2461 AGAAGACATACAGAAAGAGGGTGAATATGTGAGGCTTTTGAAATTTGTCTATTTGACA 2520  
Qy |||||  
2521 TATATGAGGTCAGATTTACTGCTTACTTGGCCACAGTGGAAACAGGAAAGAGTACATTGA 2580  
Db |||||  
2521 TATATGAGGTCAGATTTACTGCTTACTTGGCCACAGTGGAAACAGGAAAGAGTACATTGA 2580  
Qy |||||  
2581 TGAATATTTCTTGTGAGTCTGCCACCTTCTGATGGTTTGCATCTATATATGGACACA 2640  
Db |||||  
2581 TGAATATTTCTTGTGAGTCTGCCACCTTCTGATGGTTTGCATCTATATATGGACACA 2640  
Qy |||||  
2641 GAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATATGATTTGGCATTTTGTCCAGT 2700  
Db |||||  
2701 TAGATATACATTTGATGTTTGCAGTAGAAGAAATTTATCAATTTTGGCTTCAATCA 2760  
Qy |||||  
2701 TAGATATACATTTGATGTTTGCAGTAGAAGAAATTTATCAATTTTGGCTTCAATCA 2760  
Db |||||  
2761 AAGGATACACAGCCAAATATATACAGAGTGCAGAGGTTTACTAGATTTAGACA 2820  
Qy |||||  
2761 AAGGATACACAGCCAAATATATACAGAGTGCAGAGGTTTACTAGATTTAGACA 2820  
Db |||||  
2821 TGCAGACTATCAAGATAACCAAGCTAAAAATTTAAGTGGTGTGCTCAAAAAAGAAAGCTGT 2880  
Qy |||||  
2821 TGCAGACTATCAAGATAACCAAGCTAAAAATTTAAGTGGTGTGCTCAAAAAAGAAAGCTGT 2880  
Db |||||  
2881 CATTTAGAAATGCTGTTCTTGGGAACCAAGATACCTGCTGTAGATGAACCAAGCTG 2940  
Qy |||||  
2881 CATTTAGAAATGCTGTTCTTGGGAACCAAGATACCTGCTGTAGATGAACCAAGCTG 2940  
Db |||||  
2941 GAATGGACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAATACAGAAAGCCCAATC 3000  
Qy |||||  
2941 GAATGGACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAATACAGAAAGCCCAATC 3000  
Db |||||  
3001 GGGTGACAGTGTTCAGTACTCAITTCATGGATGAAGCTGACATTTCTTGAGATAGGAAAG 3060  
Qy |||||  
3001 GGGTGACAGTGTTCAGTACTCAITTCATGGATGAAGCTGACATTTCTTGAGATAGGAAAG 3060  
Db |||||  
3061 CTGTGATATACAGAGGAATGCTGAATATGTTGGTTCTTCAATGTTCTTCAAAAGTAAAT 3120  
Qy |||||  
3061 CTGTGATATACAGAGGAATGCTGAATATGTTGGTTCTTCAATGTTCTTCAAAAGTAAAT 3120  
Db |||||  
3121 GGGGGATCGGCTACCGCTCAGCATGTACATAGCAATATTTGGCCACAGATCTCTTT 3180  
Qy |||||  
3121 GGGGGATCGGCTACCGCTCAGCATGTACATAGCAATATTTGGCCACAGATCTCTTT 3180  
Db |||||  
3181 CTTTCACTGGTTAAACACATATACCTGAGCTACTTTATTAACACAGAAATGACCAAC 3240  
Qy |||||  
3181 CTTTCACTGGTTAAACACATATACCTGAGCTACTTTATTAACACAGAAATGACCAAC 3240  
Db |||||  
3241 TTGTGTATAGCTTCCCTTTCAAGGACATGGAACAAATTTTCAAGGTTTGTCTGCCCCTAG 3300  
Qy |||||  
3241 TTGTGTATAGCTTCCCTTTCAAGGACATGGAACAAATTTTCAAGGTTTGTCTGCCCCTAG 3300  
Db |||||

3301 ACAGTCATCAAAATTTGGTGTCAITTTCTTATGGTGTTCCTATGACGACTTTGGAAGACG 3360  
Db |||||  
3301 ACAGTCATCAAAATTTGGTGTCAITTTCTTATGGTGTTCCTATGACGACTTTGGAAGACG 3360  
Qy |||||  
3361 TATTTTAAAGCTAGAAAGTTGAAGCAAGAAATTTGACCAAGCAGATTTAGTGTATTTACTC 3420  
Db |||||  
3361 TATTTTAAAGCTAGAAAGTTGAAGCAAGAAATTTGACCAAGCAGATTTAGTGTATTTACTC 3420  
Qy |||||  
3421 AGCAGCCACTGGAGGAAGAAATGGATTCAAAATCTTTTGTATGAAATGGAACAGAGCTTAC 3480  
Db |||||  
3421 AGCAGCCACTGGAGGAAGAAATGGATTCAAAATCTTTTGTATGAAATGGAACAGAGCTTAC 3480  
Qy |||||  
3481 TTATTTCTTTCTGAAACCAAGGCTTCTCTAGTGAGCACCAGCTTTTGGAAACACAG 3540  
Db |||||  
3481 TTATTTCTTTCTGAAACCAAGGCTTCTCTAGTGAGCACCAGCTTTTGGAAACACAG 3540  
Qy |||||  
3541 TGTATACATAGCAAAAGTTTCATTTCTTTACCTTGAACCTGTAAGTAATCAGTGAGAT 3600  
Db |||||  
3541 TGTATACATAGCAAAAGTTTCATTTCTTTACCTTGAACCTGTAAGTAATCAGTGAGAT 3600  
Qy |||||  
3601 CAGTGTGTCTGCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTGTGTTTCATC 3660  
Db |||||  
3601 CAGTGTGTCTGCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTGTGTTTCATC 3660  
Qy |||||  
3661 ACTTTTTTAAATGCTGTGGTTCCCATCAAACTTCTTCCAGACTTATATTTTCTAAAC 3720  
Db |||||  
3661 ACTTTTTTAAATGCTGTGGTTCCCATCAAACTTCTTCCAGACTTATATTTTCTAAAC 3720  
Qy |||||  
3721 CTGGAGACAAACACATATAAATACAAACCAAGTCTGCTTCTTCAAAATTTCTGCTGACTCAG 3780  
Db |||||  
3721 CTGGAGACAAACACATATAAATACAAACCAAGTCTGCTTCTTCAAAATTTCTGCTGACTCAG 3780  
Qy |||||  
3781 ATATCAGTGCATCTTATTTAGCTTTTTCACAGCCAGAACATAATGGGACCATGATTAATG 3840  
Db |||||  
3781 ATATCAGTGCATCTTATTTAGCTTTTTCACAGCCAGAACATAATGGGACCATGATTAATG 3840  
Qy |||||  
3841 ACAGTGACATGATATCCGTGGCTCCCATAGTGGCTTTTAAATGTGATGATTCAGAAA 3900  
Db |||||  
3841 ACAGTGACATGATATCCGTGGCTCCCATAGTGGCTTTTAAATGTGATGATTCAGAAA 3900  
Qy |||||  
3901 AGGACTATGTTTGTGAGCTGTTTCAACAGTACTATGTTTATTTCTTTACCTATATTTAG 3960  
Db |||||  
3901 AGGACTATGTTTGTGAGCTGTTTCAACAGTACTATGTTTATTTCTTTACCTATATTTAG 3960  
Qy |||||  
3961 TGAATATCATTAAGTAACTACTATCTTTATCAATTTAAATGTGATGAAACCATCAGATCT 4020  
Db |||||  
3961 TGAATATCATTAAGTAACTACTATCTTTATCAATTTAAATGTGATGAAACCATCAGATCT 4020  
Qy |||||  
4021 GGAGTACCCCATCTTTTCAAGAAATTTACTGATATAGTTTAAATTTGAGCTGTATTTTC 4080  
Db |||||  
4021 GGAGTACCCCATCTTTTCAAGAAATTTACTGATATAGTTTAAATTTGAGCTGTATTTTC 4080  
Qy |||||  
4081 AAGCAGCTTGTGTGAAATCATTTGTACTGCAATGCACTTACTTTGCCATGCAAAATG 4140  
Db |||||  
4081 AAGCAGCTTGTGTGAAATCATTTGTACTGCAATGCACTTACTTTGCCATGCAAAATG 4140  
Qy |||||  
4141 CAGAGATCATTAAGATCAAGCTTACTCAACTTAACTTTTCCAGCTCTTTGCCATCTG 4200  
Db |||||  
4141 CAGAGATCATTAAGATCAAGCTTACTCAACTTAACTTTTCCAGCTCTTTGCCATCTG 4200  
Qy |||||  
4201 CATTTTGGATTTGGCAAGCTGTTGTTGATATCCCTTATTTTATCATTTTATTTTGA 4260  
Db |||||  
4201 CATTTTGGATTTGGCAAGCTGTTGTTGATATCCCTTATTTTATCATTTTATTTTGA 4260  
Qy |||||  
4261 TGTAGGAAGCTTACTGGCATTTCAATTTAGATTTATTTTATCTGTAAGTTCCTTG 4320  
Db |||||  
4261 TGTAGGAAGCTTACTGGCATTTCAATTTAGATTTATTTTATCTGTAAGTTCCTTG 4320  
Qy |||||  
4321 CTGTGTTTTTGGCTTATGTTTATGTTTCCATCAGTATTTCTGTTCACTTATATTTGCTT 4380  
Db |||||  
4321 CTGTGTTTTTGGCTTATGTTTATGTTTCCATCAGTATTTCTGTTCACTTATATTTGCTT 4380  
Qy |||||  
4381 CTTTCCACTTTTAAAGAAAATTTTAAATACCAAGAAATTTTGGTCAATTTATCTATCTGTGG 4440

Db 4381 |||||CTTTACCTTTAAAGAAATTTTAAATACCAAGAAATTTGGTCATTTATCTGTGG 4440  
Qy 4441 CAGGTTGNCCTTGTATGCAATCACTGAAATATCTTTTATGGGATACACAATTCGAA 4500  
Db 4441 CAGGTTGNCCTTGTATGCAATCACTGAAATATCTTTTATGGGATACACAATTCGAA 4500  
Qy 4501 CTATCTTCATTAATGCGCTTTGTATCATCTCAATCTCACTCTTCTAGGTTGCTGA 4560  
Db 4501 CTATCTTCATTAATGCGCTTTGTATCATCTCAATCTCACTCTTCTAGGTTGCTGA 4560  
Qy 4561 TTTCTTTTCAAAAGATTTCTTGAAGAAATGTAAGAAATATGGAACCTTATATCCAT 4620  
Db 4561 TTTCTTTTCAAAAGATTTCTTGAAGAAATGTAAGAAATATGGAACCTTATATCCAT 4620  
Qy 4621 GGGATAGGCTTTCAGTAGCTTTATATCGCTTTACCTGCGAGTGTACTGTGGATTTTCC 4680  
Db 4621 GGGATAGGCTTTCAGTAGCTTTATATCGCTTTACCTGCGAGTGTACTGTGGATTTTCC 4680  
Qy 4681 TCTTACAAATCTATGAGAAATATATGAGGCGAGATCAATAGAAAGATCCCTTTTCA 4740  
Db 4681 TCTTACAAATCTATGAGAAATATATGAGGCGAGATCAATAGAAAGATCCCTTTTCA 4740  
Qy 4741 GAAACCTTTCAACGAGCTTAAATATAGGAAGCTTCCAGAACCCAGACAATGAGGATG 4800  
Db 4741 GAAACCTTTCAACGAGCTTAAATATAGGAAGCTTCCAGAACCCAGACAATGAGGATG 4800  
Qy 4801 AAGATGAAGTGTCAAAAGCTGAAAGCTAAAGGCTAAAGGCTGATGGGTGCGAGTGT 4860  
Db 4801 AAGATGAAGTGTCAAAAGCTGAAAGCTAAAGGCTAAAGGCTGATGGGTGCGAGTGT 4860  
Qy 4861 GTGAGGAGAACCATTCATTTATGTCAGCAATTTGCAATGAAGATATGATGACAAGAAAG 4920  
Db 4861 GTGAGGAGAACCATTCATTTATGTCAGCAATTTGCAATGAAGATATGATGACAAGAAAG 4920  
Qy 4921 ATTTCTTCTTTCAAGAAAGTAAAGAAAGTGGCAACTAATACATCTCTTCTGTGTA 4980  
Db 4921 ATTTCTTCTTTCAAGAAAGTAAAGAAAGTGGCAACTAATACATCTCTTCTGTGTA 4980  
Qy 4981 AAAAGAGAGATCTTAGGACTATTTGGTCCAAATGGTGTGCAAGCAACAATTTATTA 5040  
Db 4981 AAAAGAGAGATCTTAGGACTATTTGGTCCAAATGGTGTGCAAGCAACAATTTATTA 5040  
Qy 5041 ATATCTGTTGTTGATTAATTAACCAACTTCAGCCAGGATTTTATAGGAGATTTCTT 5100  
Db 5041 ATATCTGTTGTTGATTAATTAACCAACTTCAGCCAGGATTTTATAGGAGATTTCTT 5100  
Qy 5101 CAGAGACAAGTGAAGATGATGATTTCACTGAAGTGTATGGGTTACTGTCTCAGATAACC 5160  
Db 5101 CAGAGACAAGTGAAGATGATGATTTCACTGAAGTGTATGGGTTACTGTCTCAGATAACC 5160  
Qy 5161 CTTTGTGGCCAGATACATACATTTGAGGAAATTTTGAATTTATGGAGCTGTCAAGGAA 5220  
Db 5161 CTTTGTGGCCAGATACATACATTTGAGGAAATTTTGAATTTATGGAGCTGTCAAGGAA 5220  
Qy 5221 TGAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5280  
Db 5221 TGAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5280  
Qy 5281 AACATCTTCAGAGACTGTAAAGAAATACCTGCGAGGAAATCAACAGCACTGATTTAAAG 5340  
Db 5281 AACATCTTCAGAGACTGTAAAGAAATACCTGCGAGGAAATCAACAGCACTGATTTAAAG 5340  
Qy 5341 CTCTAAGTATGCTAGGAAATCTCAGATTTCTTGTAGATGAACCATCTACAGGATG 5400  
Db 5341 CTCTAAGTATGCTAGGAAATCTCAGATTTCTTGTAGATGAACCATCTACAGGATG 5400  
Qy 5401 ATCCAAAGCCAAACAGCATGTGGCGAGCAATTCGAACTGATTTAAAGCAGAAAGC 5460  
Db 5401 ATCCAAAGCCAAACAGCATGTGGCGAGCAATTCGAACTGATTTAAAGCAGAAAGC 5460  
Qy 5461 GGGCTGCTATTTGACCACTCACTATATGAGGAGGCGAGGCTGTCTGTGATCGAGTAG 5520

Db 5461 GGGCTGCTATTTCTGACCACTCACTATATGAGGAGGCGAGAGGCTGTCTGTGATCGAGTAG 5520  
Qy 5521 CTATCATGCTGTCTGGGCGAGTTAAGATGTATCGGAACTAGTACAACTATAAGAGTAAAT 5580  
Db 5521 CTATCATGCTGTCTGGGCGAGTTAAGATGTATCGGAACTAGTACAACTATAAGAGTAAAT 5580  
Qy 5581 TTGGAAGAGGCTACTTTTGGAAATTTAAATTTGAAAGCTGATGAGAAACCTTAGAAGTAG 5640  
Db 5581 TTGGAAGAGGCTACTTTTGGAAATTTAAATTTGAAAGCTGATGAGAAACCTTAGAAGTAG 5640  
Qy 5641 ACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCGCTCAGGAAAGTTTTT 5700  
Db 5641 ACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCGCTCAGGAAAGTTTTT 5700  
Qy 5701 CTTCTATTTTGGCTTATATAATTTCTTAAGGAAGATGTTCACTCCCTTTTCAATCTTTT 5760  
Db 5701 CTTCTATTTTGGCTTATATAATTTCTTAAGGAAGATGTTCACTCCCTTTTCAATCTTTT 5760  
Qy 5761 TTAAGCTGGAAGAGCTAAACATGCTTTTGCATTTGAAGAAATATAGCTTTTCTCAAGCAA 5820  
Db 5761 TTAAGCTGGAAGAGCTAAACATGCTTTTGCATTTGAAGAAATATAGCTTTTCTCAAGCAA 5820  
Qy 5821 CATTTGGAACAGGTTTGTAGAACTCACTAAAGAACCAAGAGAGGAGAGATATAGTTGTG 5880  
Db 5821 CATTTGGAACAGGTTTGTAGAACTCACTAAAGAACCAAGAGAGGAGAGATATAGTTGTG 5880  
Qy 5881 GAACTTTAAACAGCACACTTTTGGTGGGAAACCAACAAAGAGATAGAGTAGTATTTTCAA 5940  
Db 5881 GAACTTTAAACAGCACACTTTTGGTGGGAAACCAACAAAGAGATAGAGTAGTATTTTCAA 5940  
Qy 5941 TTTGTATTTTGGTCTGCTTACTTGGGACTTCTTTTCACTTAAATTTTAACTTTTGG 6000  
Db 5941 TTTGTATTTTGGTCTGCTTACTTGGGACTTCTTTTCACTTAAATTTTAACTTTTGG 6000  
Qy 6001 TTTAAAGAGTTTTTATTTGGAATGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6060  
Db 6001 TTTAAAGAGTTTTTATTTGGAATGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6060  
Qy 6061 TAAGTCTCTTAATGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6120  
Db 6061 TAAGTCTCTTAATGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6120  
Qy 6121 GTATAATTAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6180  
Db 6121 GTATAATTAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6180  
Qy 6181 TCATCTTTTTCACCACTCAGAAACAGTCTTCTGAAATTTGTAATTTAAAGGAAATTTGTA 6240  
Db 6181 TCATCTTTTTCACCACTCAGAAACAGTCTTCTGAAATTTGTAATTTAAAGGAAATTTGTA 6240  
Qy 6241 AGAATAGTTTTTATTTTAAAGTTATCTTTTAAAGTTATGCTTCTTAAATTAAGTACGTA 6300  
Db 6241 AGAATAGTTTTTATTTTAAAGTTATCTTTTAAAGTTATGCTTCTTAAATTAAGTACGTA 6300  
Qy 6301 ATGTTTCCAAATCTAAATAAAAACTAAATACTAATGCTAGAGAAAGATACATAAAGC 6360  
Db 6301 ATGTTTCCAAATCTAAATAAAAACTAAATACTAATGCTAGAGAAAGATACATAAAGC 6360  
Qy 6361 AATGTGAAGTTTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 6420  
Db 6361 AATGTGAAGTTTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 6420  
Qy 6421 TGTCTATCCGTAAGAGCTGAAGTGAAGCACTAGGAAATCTCAATATAGAGATTTGAGGAA 6480  
Db 6421 TGTCTATCCGTAAGAGCTGAAGTGAAGCACTAGGAAATCTCAATATAGAGATTTGAGGAA 6480  
Qy 6481 AGTTATATCCACTAGTGTGCGAGTCAATGATCAATAAGTGAAT 6525  
Db 6481 AGTTATATCCACTAGTGTGCGAGTCAATGATCAATAAGTGAAT 6525

```
; Sequence 4, Application US/10090458
; Publication No. US20020123107A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hongyun
; APPLICANT: Kilinski, Ligia
; APPLICANT: Le Bihan, Stephane
; TITLE OF INVENTION: NOVEL ABCA5 TRANSPORTER AND USES THEREOF
; FILE REFERENCE: 100103.403
; CURRENT APPLICATION NUMBER: US/10/090.458
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 5475
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 25
; OTHER INFORMATION: n = A,T,C or G
US-10-090-458-4

Query Match      82.3%; Score 5373.2; DB 14; Length 5475;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5375; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 994 AGGTTATTTCAGAAAAATGTCACATGTCACATGCAATTAGGGAGGTAGGAGTTTGGAGACAGACCA 1053
Db 97 AGGTTATTTCAGAAAAATGTCACATGTCACATTAGGGAGGTAGGAGTTTGGAGACAGACCA 156

Qy 1054 GAACACTTCTACTGAAGAATTACTTAATAATGCGAGAACCAAAAAAGAGTAGTGTTCAGG 1113
Db 157 GAACACTTCTACTGAAGAATTACTTAATAATGCGAGAACCAAAAAAGAGTAGTGTTCAGG 216

Qy 1114 AAATCTTTTCCACTATTTTTTTTATTTTGGTTAATTAATAGCATGATGCATCCAA 1173
Db 217 AAATCTTTTCCACTATTTTTTTTATTTTGGTTAATTAATAGCATGATGCATCCAA 276

Qy 1174 ATAAGAAATAGGAAGTGCCTAATATAGAACTCAATCTCTATGGCAAGTTTACTCTTT 1233
Db 277 ATAAGAAATAGGAAGTGCCTAATATAGAACTCAATCTCTATGGCAAGTTTACTCTTT 336

Qy 1234 CTAATCTAATTTTGGATATCTCCAGTGACTAATATTAACAAGCAGCATCATCGAGAAAG 1293
Db 337 CTAATCTAATTTCTGGATATCTCCAGTGACTAATATTAACAAGCAGCATCATCGAGAAAG 396

Qy 1294 TGTCTACTGATCATCTACCTGATGTCTAATATTACTGAAGATATACAATGAAAAAGAA 1353
Db 397 TGTCTACTGATCATCTACCTGATGTCTAATATTACTGAAGATATACAATGAAAAAGAA 456

Qy 1354 TGTTAAATCCAGTCTCTTAAGCCGAGCAACTTTGTAGGTGTGTTTCAAAGACTCCA 1413
Db 457 TGTTAAATCCAGTCTCTTAAGCCGAGCAACTTTGTAGGTGTGTTTCAAAGACTCCA 516

Qy 1414 TGTCTATGAACCTCGTTTTTTTCTGATATGATTCAGATATCTTCTATTATATGGAAT 1473
Db 517 TGTCTATGAACCTCGTTTTTTTCTGATATGATTCAGATATCTTCTATTATATGGAAT 576

Qy 1474 CAAGAGCTGCTGTTTCAAAATCATGTGAGGCTGCTAGTACTGCTCAGGTTTCACAG 1533
Db 577 CAAGAGCTGCTGTTTCAAAATCATGTGAGGCTGCTAGTACTGCTCAGGTTTCACAG 636

Qy 1534 TTTTAAAGCATCATAGATGCTGCCATTATACAGTTTGAAGACCAATGTTTCTCTTTGGA 1593
Db 637 TTTTAAAGCATCATAGATGCTGCCATTATACAGTTTGAAGACCAATGTTTCTCTTTGGA 696

Qy 1594 AGGAGCTGGAGTCAACTAAAGCTGTTTATTAATGGAGAAACTGCTGTTGTAGAAATAGATA 1653
Db 697 AGGAGCTGGAGTCAACTAAAGCTGTTTATTAATGGAGAAACTGCTGTTGTAGAAATAGATA 756

Qy 1654 CCTTTCCCGAGGAGTAATTTTAAATATACCTAGTTATAGCAATTTTCACTTTTGGATCT 1713
Db 757 CCTTTCCCGAGGAGTAATTTTAAATATACCTAGTTATAGCAATTTTCACTTTTGGATCT 816

1714 TTTTGGCAATTCATATCGTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAA 1773
817 TTTTGGCAATTCATATCGTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAA 876
1774 TGGGACTTCATGATACTGCTTTTGGCTTTCCTGGGTTCTTCTATATACAAAGTTTAAATTT 1833
877 TGGGACTTCATGATACTGCTTTTGGCTTTCCTGGGTTCTTCTATATACAAAGTTTAAATTT 936
1834 TTTTATGTCCTTCTTATGCGGAGTCATTCGACAGCTTCTTTGTTTATTTCCCTCAAAGTA 1893
937 TTTTATGTCCTTCTTATGCGGAGTCATTCGACAGCTTCTTTGTTTATTTTCCCTCAAAGTA 996
1894 GCAGCAATTTGATATTTCTGCTTTTCTTTTATGGATTTATCATCTGATATTTTTCGCTT 1953
997 GCAGCAATTTGATATTTCTGCTTTTCTTTTATGGATTTATCATCTGATATTTTTCGCTT 1056
1954 TAAATGCTGACACCTCTTTTAAATAATCAAAACATGTGGGAATAGTTGAAATTTTGTGTTA 2013
1057 TAAATGCTGACACCTCTTTTAAATAATCAAAACATGTGGGAATAGTTGAAATTTTGTGTTA 1116
2014 CTGTGGCTTTTGGATTTATTTGGCTTTATGATATCTCATAGAAAGTTTTCCTCAAAATCGT 2073
1117 CTGTGGCTTTTGGATTTATTTGGCTTTATGATATCTCATAGAAAGTTTTCCTCAAAATCGT 1176
2074 TAGTGTGGCTTTTTCAGTCTCTTCTGTCACTGTACTTTTGTGATTTGTTGACAGGTCA 2133
1177 TAGTGTGGCTTTTTCAGTCTCTTCTGTCACTGTACTTTTGTGATTTGTTGACAGGTCA 1236
2134 TGCATTTAGAGATTTTAAATGAAGTGTCTTCAATTTTCAATTTGACTCGAGGCCCATAATC 2193
1237 TGCATTTAGAGATTTTAAATGAAGTGTCTTCAATTTTCAATTTTGAATGCTGCGAGGCCATATC 1296
2194 CTCTAATTAATCAATTTATCATCTCACACTTAATAGTATATCTATGCTCTCTGCTGCTG 2253
1297 CTCTAATTAATCAATTTATCATCTCACACTTAATAGTATATTTCTATGCTCTCTGCTGCTG 1356
2254 TCTATCTTGATCAAGTCAATTTCCAGGGGAATTTGGCTTTACGGAGATCATCTTTATATTTTC 2313
1357 TCTATCTTGATCAAGTCAATTTCCAGGGGAATTTGGCTTTACGGAGATCATCTTTATATTTTC 1416
2314 TGAAGCTTCAATATTTGGTCAAGAGTAAAGAAATTTATGAGGAGTTTATCAGAGGGCAATG 2373
1417 TGAAGCTTCAATATTTGGTCAAGAGTAAAGAAATTTATGAGGAGTTTATCAGAGGGCAATG 1476
2374 TTAATGGAAATATTTAGTTTGTGAAATTTATGAGCCAGTTTCTTCTGAGAAATTTGTAGGAA 2433
1477 TTAATGGAAATATTTAGTTTGTGAAATTTATGAGCCAGTTTCTTCTGAGAAATTTGTAGGAA 1536
2434 AAGAAGCCATAAGAAATTTAGTGGTATTTCAAGAGACATACAGAAAGAGGGTGAATAATGTGG 2493
1537 AAGAAGCCATAAGAAATTTAGTGGTATTTCAAGAGACATACAGAAAGAGGGTGAATAATGTGG 1596
2494 AGGCTTTGAGAAATTTGTCAATTTGACATATATGAGGGTCAAGATTAAGTCTTACTTGGCC 2553
1597 AGGCTTTGAGAAATTTGTCAATTTGACATATATGAGGGTCAAGATTAAGTCTTACTTGGCC 1656
2554 ACAGTGGAGAGGAGAGTACATTTGATCAATATTTCTTTGTGAGCTCTGCCCCACCTTCTG 2613
1657 ACAGTGGAGAGGAGAGTACATTTGATCAATATTTCTTTGTGAGCTCTGCCCCACCTTCTG 1716
2614 ATGGGTTTTCATCTATATATGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAA 2673
1717 ATGGGTTTTCATCTATATATGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAA 1776
2674 GAAAAATGATGGCAATTTGTCCACAGTTAGATATACATTTGATGTTTGTGACAGTGAAG 2733
1777 GAAAAATGATGGCAATTTGTCCACAGTTAGATATACATTTGATGTTTGTGACAGTGAAG 1836
2734 AAAAAATTAATTTGGCTTCAATCAAAAGGAGTACACGCAACAATATAAACAAGAG 2793
1837 AAAAAATTAATTTGGCTTCAATCAAAAGGAGTACACGCAACAATATAAACAAGAG 1896
```

Qy	2794	TGCAGAGGTTT	TACTAGATTT	TAGACATG	CAGACTAT	CAAAAGATA	CAACAGCT	TAAAAAT	2855
Db	1897	TGCAGAGGTTT	TACTAGATTT	TAGACATG	CAGACTAT	CAAAAGATA	CAACAGCT	TAAAAAT	1956
Qy	2854	TAAGTGTGGT	CAAAAAGAAC	GTGCATTT	TAGGAAT	TGCTGTT	TCTTTGG	GAACCCAAAGA	2913
Db	1957	TAAGTGTGGT	CAAAAAGAAC	GTGCATTT	TAGGAAT	TGCTGTT	TCTTTGG	GAACCCAAAGA	2016
Qy	2914	TACTGCTGCT	TAGATGA	ACCAACAG	CTGGAA	TGGACCC	CTGTTTC	CGACATATTG	2973
Db	2017	TACTGCTGCT	TAGATGA	ACCAACAG	CTGGAA	TGGACCC	CTGTTTC	CGACATATTG	2076
Qy	2974	ATCTTTTAAA	ATA	CAGAAAAC	CCAATCG	GGTGA	CAGTGT	TCAGTACTCAT	3033
Db	2077	ATCTTTTAAA	ATA	CAGAAAAC	CCAATCG	GGTGA	CAGTGT	TCAGTACTCAT	2136
Qy	3034	AAGCTGACAT	TTTGC	AGATAG	GAAGAA	AGCTGT	GCATAT	CACAAGGAAT	3093
Db	2137	AAGCTGACAT	TTTGC	AGATAG	GAAGAA	AGCTGT	GCATAT	CACAAGGAAT	2196
Qy	3094	GTTCTTCAAT	GTTCCT	CAAAAAG	TAAATG	GGGGG	ATCGGCT	ACCGCTC	3153
Db	2197	GTTCTTCAAT	GTTCCT	CAAAAAG	TAAATG	GGGGG	ATCGGCT	ACCGCTC	2256
Qy	3154	ACAAATAT	TG	CCACAGAA	TCTCTTT	CTTCA	CTCGT	TGTTAAACAA	3213
Db	2257	ACAAATAT	TG	CCACAGAA	TCTCTTT	CTTCA	CTCGT	TGTTAAACAA	2316
Qy	3214	CTTTTATTA	CAACAGAA	TGACCA	ACAACTT	GTGTAT	AGCTTT	CAAGGACAT	3273
Db	2317	CTTTTATTA	CAACAGAA	TGACCA	ACAACTT	GTGTAT	AGCTTT	CAAGGACAT	2376
Qy	3274	AATTTTCA	G	GGTTGTTT	CTGCCCT	TAGACAG	CTCAIT	CAAAATTT	3333
Db	2377	AATTTTCA	G	GGTTGTTT	CTGCCCT	TAGACAG	CTCAIT	CAAAATTT	2436
Qy	3334	GTGTTTCC	ATGACGACT	TTTGA	AGACG	TATTTT	AAAGCT	ATAGAAGTT	3393
Db	2437	GTGTTTCC	ATGACGACT	TTTGA	AGACG	TATTTT	AAAGCT	ATAGAAGTT	2496
Qy	3394	ACCAAGCAG	ATATAG	TGTATTT	TACTC	AGCAGC	CACTTGG	AGGAGAGAA	3453
Db	2497	ACCAAGCAG	ATATAG	TGTATTT	TACTC	AGCAGC	CACTTGG	AGGAGAGAA	2556
Qy	3454	CTTTTGTAG	AAATGG	AAACAGAG	CTTACT	TATTTCT	TGAAA	CAAGGCTT	3513
Db	2557	CTTTTGTAG	AAATGG	AAACAGAG	CTTACT	TATTTCT	TGAAA	CAAGGCTT	2616
Qy	3514	GCACCATG	AGCCTTT	TGGAAA	CAACAGAT	TGTATA	CAATAG	CAGAAAGTT	3573
Db	2617	GCACCATG	AGCCTTT	TGGAAA	CAACAGAT	TGTATA	CAATAG	CAGAAAGTT	2676
Qy	3574	TGAAAACGT	GAAAGTAA	ATCAGT	GAGATCAG	TGTTGCTT	CTGCTTTT	TAATTTTTT	3633
Db	2677	TGAAAACGT	GAAAGTAA	ATCAGT	GAGATCAG	TGTTGCTT	CTGCTTTT	TAATTTTTT	2736
Qy	3634	TTCCAGAT	TTTTTAT	GTTTTT	TGGTTCAT	CACTCT	TTTTT	AAAAATG	3693
Db	2737	TTCCAGAT	TTTTTAT	GTTTTT	TGGTTCAT	CACTCT	TTTTT	AAAAATG	2796
Qy	3694	TTGTTTCC	AGACTTAT	TTTTT	CTAAAA	CCCTGGAG	CAAAACCA	CATATAA	3753
Db	2797	TTGTTTCC	AGACTTAT	TTTTT	CTAAAA	CCCTGGAG	CAAAACCA	CATATAA	2856
Qy	3754	TGCTTCTT	CAAAAAT	TTCTGCT	GACTCAG	ATATCAG	TGATCTT	TATTTAG	3813
Db	2857	TGCTTCTT	CAAAAAT	TTCTGCT	GACTCAG	ATATCAG	TGATCTT	TATTTAG	2916
Qy	3814	AGAAATAA	TGCTG	CAGATGA	TAAATG	ACGTG	ACTATG	TATCCG	3873
Db	2917	AGAAATAA	TGCTG	CAGATGA	TAAATG	ACGTG	ACTATG	TATCCG	2976
Qy	3874	CGGCTTTAA	T	GTGATG	CAATTT	CAGAAA	AGGACTAT	TGTTTTT	3933

Db	2977	CGGCTTTAAAGTGATGCATTCAGAAAAGCATATGTTTTGCGAGCTGTTTTCAACAGTA	3036
Qy	3934	CTATGGTTTATCTTTTACCTATATTAGTGAATACATTAAGTAACACTACTCTTTATCATTT	3993
Db	3037	CTATGGTTTATCTTTTACCTATATTAGTGAATATCATTAAGTAACACTACTCTTTATCATTT	3096
Qy	3994	TAAATGTGACTGAACCAATCCAGATCTGGAGTACCCCATCTCTTTCAAGAAAATTTACTTGATA	4053
Db	3097	TAAATGTGACTGAACCAATCCAGATCTGGAGTACCCCATCTCTTTCAAGAAAATTTACTTGATA	3156
Qy	4054	TAGTTTTTAAAAATTGAGCTGTATTTTCAAGCAGCTTTTGGTGGAAATCATTTGTTTACTTGCAA	4113
Db	3157	TAGTTTTTAAAAATTGAGCTGTATTTTCAAGCAGCTTTTGGTGGAAATCATTTGTTTACTTGCAA	3216
Qy	4114	TGCCACCTTACTTTGGCCATGGAAAATCCAGAGAATCATTAAGATCAAAAGCTTTACTCAAC	4173
Db	3217	TGCCACCTTACTTTGGCCATGGAAAATCCAGAGAATCATTAAGATCAAAAGCTTTACTCAAC	3276
Qy	4174	TTAAACCTTTCAGGCTTTTGGCCATCTGCATATTGGATTTGGAAGCTGTTGTTTGTGATATCC	4233
Db	3277	TTAAACCTTTCAGGCTTTTGGCCATCTGCATATTGGATTTGGAAGCTGTTGTTTGTGATATCC	3336
Qy	4234	CCCTATTTTTTATCATTTCTTATTTTGGATCTGAGAGCTTTACTTGGCAATTTTCATATGGAT	4293
Db	3337	CCCTATTTTTTATCATTTCTTATTTTGGATCTGAGAGCTTTATTTGGCAATTTTCATATGGAT	3396
Qy	4294	TATATTTTTTATCTGTAAAGTTCCTTCTGTGTGTTTTTGGCCCTATTATGGTTATGTTCCAT	4353
Db	3397	TATATTTTTTATCTGTAAAGTTCCTTCTGTGTGTTTTTGGCCCTATTATGGTTATGTTCCAT	3456
Qy	4354	CAGTTATTCGTTCACCTATPATTTGCTTCTTTTACCCTTAAAGAAAATTTTAAATACCAAG	4413
Db	3457	CAGTTATTCGTTCACCTATPATTTGCTTCTTTTACCCTTAAAGAAAATTTTAAATACCAAG	3516
Qy	4414	AAATTTGGCTCATTTATCTATTTCTGTGGCAGCGTTGCTGTATTGCAATCACCTGGAATAA	4473
Db	3517	AAATTTGGCTCATTTATCTATTTCTGTGGCAGCGTTGCTGTATTGCAATCACCTGGAATAA	3576
Qy	4474	CTTTCTTTATGGGATACACAATTTGCAACTATTTCTTCATTTATGCTTTTGTGATCATCATTC	4533
Db	3577	CTTTCTTTATGGGATACACAATTTGCAACTATTTCTTCATTTATGCTTTTGTGATCATCATTC	3636
Qy	4534	CAATCTATCCACTTCTAGGTTGCCTGATTTCTTTCAFAAAGATTTCTTGAAGAATGTATC	4593
Db	3637	CAATCTATCCACTTCTAGGTTGCCTGATTTCTTTCAFAAAGATTTCTTGAAGAATGTATC	3696
Qy	4594	GAAAAAATGGGACACCTATAAATCCATGGATAGGCTTTTCAGTAGCTGTTTATATCCGCTT	4653
Db	3697	GAAAAAATGGGACACCTATAAATCCATGGATAGGCTTTTCAGTAGCTGTTTATATCCGCTT	3756
Qy	4654	ACCTGCACTGTGTACTGTGGATTTTCTCTTCAATACTATGAGAAAAAATATGGAGGCA	4713
Db	3757	ACCTGCACTGTGTACTGTGGATTTTCTCTTCAATACTATGAGAAAAAATATGGAGGCA	3816
Qy	4714	GATCAATTAAGAAAAGATCCCTTTTTTCAGAAACCTTTCAACGAAGTCTAAAAATAGGAAGC	4773
Db	3817	GATCAATTAAGAAAAGATCCCTTTTTTCAGAAACCTTTCAACGAAGTCTAAAAATAGGAAGC	3876
Qy	4774	TTCCAGAACCAACAGACCAATGAGAGTGAAGATGTCAAGCTGGAAGACTTAAGG	4833
Db	3877	TTCCAGAACCAACAGACCAATGAGAGTGAAGATGTCAAGCTGGAAGACTTAAGG	3936
Qy	4834	TCAAGAGCTGTATGGGTTGCGAGTTGTCAGGAGAAACCATCCATTTATGCTCAGCAATTT	4893
Db	3937	TCAAGAGCTGTATGGGTTGCGAGTTGTCAGGAGAAACCATCCATTTATGCTCAGCAATTT	3996
Qy	4894	TGCATAAAGAAATATGATGACAAGAAAATTTTTCTTTTCAAGAAAAGTAAAGAAAGTGG	4953
Db	3997	TGCATAAAGAAATATGATGACAAGAAAAGTTTTTCTTTCTTTTCAAGAAAAGTAAAGAAAGTGG	4056
Qy	4954	CAACTAAATACATCTCTTTCTGTGTGAAAAAAGGAGAGATCTTAGGACTATTTGGGTCCAA	5013



1354 TGTTAAATCCAGTCTCTTAAGCCGAGCAACTTTGTAGGTGTGGTTTTCAGAGACTCCA 1413  
Db  
457 TGTTAAATCCAGTCTCTTAAGCCGAGCAACTTTGTAGGTGTGGTTTTCAGAGACTCCA 516  
Qy  
1414 TGTCTATGAACCTTGTGTTTTTTCCTGATATGATTCAGATATCTTCTATTTATATGATTT 1473  
Db  
517 TGTCTATGAACCTTGTGTTTTTTCCTGATATGATTCAGATATCTTCTATTTATATGATTT 576  
Qy  
1474 CAAGAGCTGGCTGTTCAAAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGGTTTCACAG 1533  
Db  
577 CAAGAGCTGGCTGTTCAAAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGGTTTCACAG 636  
Qy  
1534 TTTTACAAGCATCATATGATGCTGCCAATATACAGTTGAAGCAATGTTTCTCTTTTGA 1593  
Db  
637 TTTTACAAGCATCATATGATGCTGCCAATATACAGTTGAAGCAATGTTTCTCTTTGA 696  
Qy  
1594 AGGAGCTGGAGTCAACTAAAGCTGTTTATATGGAGAAACTGCTGTGTAGAAATAGATA 1653  
Db  
697 AGGAGCTGGAGTCAACTAAAGCTGTTTATATGGAGAAACTGCTGTGTAGAAATAGATA 756  
Qy  
1654 CCTTTCCCGAGAGTAAATTTAATATACCTAGTTATAGCAATTTTACCTTTTGGATACT 1713  
Db  
757 CCTTTCCCGAGAGTAAATTTAATATACCTAGTTATAGCAATTTTACCTTTTGGATACT 816  
Qy  
1714 TTTTGGCAATTCATATCGTAGCAAGAAAGAAAGAAATTTTAAAGATAA 1773  
Db  
817 TTTTGGCAATTCATATCGTAGCAAGAAAGAAAGAAATTTTAAAGATAA 876  
Qy  
1774 TGGAGCTTCATGATPACCTTTTGGCTTTTCTGGGTTCTTCTATATACAAGTTTAAATTT 1833  
Db  
877 TGGAGCTTCATGATPACCTTTTGGCTTTTCTGGGTTCTTCTATATACAAGTTTAAATTT 936  
Qy  
1834 TTCTTATGTCCTTTTATGGAGTCAATTCGGACAGCTTCTTTGTTATTTCTCCTCAAAGTA 1893  
Db  
937 TTCTTATGTCCTTTTATGGAGTCAATTCGGACAGCTTCTTTGTTATTTCTCCTCAAAGTA 996  
Qy  
1894 GCAGCATTTGATATTTCTGCTTTTCTTTTATGATTTATCATCTGATTTTGTGCTT 1953  
Db  
997 GCAGCATTTGATATTTCTGCTTTTCTTTTATGATTTATCATCTGATTTTGTGCTT 1056  
Qy  
1954 TAATGCTGACACCTCTTTTAAAGAAATCAAAACATGTGGAAATAGTTGAAATTTTGTGA 2013  
Db  
1057 TAATGCTGACACCTCTTTTAAAGAAATCAAAACATGTGGAAATAGTTGAAATTTTGTGA 1116  
Qy  
2014 CTGTGGCTTTTGGATTTATTTGGCTTTATGATATTCCTCATAGAAAGTTTCCCAATCGT 2073  
Db  
1117 CTGTGGCTTTTGGATTTATTTGGCTTTATGATATTCCTCATAGAAAGTTTCCCAATCGT 1176  
Qy  
2074 TAGTGTGGCTTTTTCAGTCTTTCTGTCACGTACTTTTGTGATTTGTCACAGTCA 2133  
Db  
1177 TAGTGTGGCTTTTTCAGTCTTTCTGTCACGTACTTTTGTGATTTGTCACAGTCA 1236  
Qy  
2134 TGCAATTTAGAGATTTTAAATGAAGTCTTCAATTTCAAAATTTGACTGCGGCCCATATC 2193  
Db  
1237 TGCAATTTAGAGATTTTAAATGAAGTCTTCAATTTCAAAATTTGACTGCGGCCCATATC 1296  
Qy  
2194 CTCTAATTTATCAATTTATCATGCTCAGCTTAAATAGTATTTCTATGCTCTTGGCTG 2253  
Db  
1297 CTCTAATTTATCAATTTATCATGCTCAGCTTAAATAGTATTTCTATGCTCTTGGCTG 1356  
Qy  
2254 TCTATCTTGATCAAGTCAATTTCCAGGGAAATTTGGCTTTACGGAGATCATCTTTATATTTTC 2313  
Db  
1357 TCTATCTTGATCAAGTCAATTTCCAGGGAAATTTGGCTTTACGGAGATCATCTTTATATTTTC 1416  
Qy  
2314 TGAAGCTTTCATATTTGCTCAAGAGTAAAGAAATTTATGAGGAGTTTATCAGAGGGCAATG 2373  
Db  
1417 TGAAGCTTTCATATTTGCTCAAGAGCAAAAGAAATTTATGAGGAGTTTATCAGAGGGCAATG 1476  
Qy  
2374 TTAATGAAATATTTAGTTTTTATGAAATTTATGAGCCAGTTTCTTCAAGAAATTTGTAGGAA 2433  
Db  
1477 TTAATGAAATATTTAGTTTTTATGAAATTTATGAGCCAGTTTCTTCAAGAAATTTGTAGGAA 1536

2434 AAGAGCCCATAGAAATTTAGTGGTATTTCAAGAGACATACAGAAAGAGGGTGAAATGTGG 2493  
Db  
1537 AAGAGCCCATAGAAATTTAGTGGTATTTCAAGAGACATACAGAAAGAGGGTGAAATGTGG 1596  
Qy  
2494 AGGCTTTGAGAAATTTGTCAATTTGACATATATGAGGGTCAGATTACTGCTTACTTGGCC 2553  
Db  
1597 AGGCTTTGAGAAATTTGTCAATTTGACATATATGAGGGTCAGATTACTGCTTACTTGGCC 1656  
Qy  
2554 ACAGTGGAAACAGGAAGAGTAGTACATTTGATGAATATTTCTTTGTGGACTCTGCCCACCTCTG 2613  
Db  
1657 ACAGTGGAAACAGGAAGAGTAGTACATTTGATGAATATTTCTTTGTGGACTCTGCCCACCTCTG 1716  
Qy  
2614 ATGGCTTTGCATCTATATATATGACACAGAGTCTCAGAAATAGATGAATTTTGAAGCAA 2673  
Db  
1717 ATGGCTTTGCATCTATATATGACACAGAGTCTCAGAAATAGATGAATTTTGAAGCAA 1776  
Qy  
2674 GAAATATGATTTGGCATTTTGTCCACAGTTAGATATACATTTGATGTTTTTGACAGTAGAAG 2733  
Db  
1777 GAAATATGATTTGGCATTTTGTCCACAGTTAGATATACATTTGATGTTTTTGACAGTAGAAG 1836  
Qy  
2734 AAAATTTTATCAATTTTGGCTTTCAATCAAGGGATACACAGCCAAACAATATATACAAGAG 2793  
Db  
1837 AAAATTTTATCAATTTTGGCTTTCAATCAAGGGATACACAGCCAAACAATATATACAAGAG 1896  
Qy  
2794 TGCAGAGGTTTTTACTAGATTTAGACATGACAGCTTATCAAGATTAACAAGCTAAAAAT 2853  
Db  
1897 TGCAGAGGTTTTTACTAGATTTAGACATGACAGCTTATCAAGATTAACAAGCTAAAAAT 1956  
Qy  
2854 TAAAGTGGTTCAGAAAGAAAGCTGCTATTAGGAATTTGCTTTGGGAAACCCAAAGA 2913  
Db  
1957 TAAAGTGGTTCAGAAAGAAAGCTGCTATTAGGAATTTGCTTTGGGAAACCCAAAGA 2016  
Qy  
2914 TACTGCTCTAGATGAACCAACAGCTGGAATGGAGCCCTGTTCTCGACATATTTGATGGA 2973  
Db  
2017 TACTGCTCTAGATGAACCAACAGCTGGAATGGAGCCCTGTTCTCGACATATTTGATGGA 2076  
Qy  
2974 ATCTTTTAAAAATACAGAAAAAGCAATCGGGTCACAGTGTTCAGTACTCATTTTCATGGATG 3033  
Db  
2077 ATCTTTTAAAAATACAGAAAAAGCAATCGGGTCACAGTGTTCAGTACTCATTTTCATGGATG 2136  
Qy  
3034 AAGCTGACATTTCTGACATAGGAAGCTGTGATATCAAGAGGATGCTGAAATGTGTTG 3093  
Db  
2137 AAGCTGACATTTCTGACATAGGAAGCTGTGATATCAAGAGGATGCTGAAATGTGTTG 2196  
Qy  
3094 GTTCTTCAATGTTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 3153  
Db  
2197 GTTCTTCAATGTTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2256  
Qy  
3154 ACAATATTTGTGCCACAGAAATCTCTTTCTTCACTGGTTTAAACAACATATACCTGGAGCTA 3213  
Db  
2257 ACAATATTTGTGCCACAGAAATCTCTTTCTTCACTGGTTTAAACAACATATACCTGGAGCTA 2316  
Qy  
3214 CTTTATTAACAAGAAATGACCAACTGTGTATAGCTTGGCTTTCAAGGACATGGACA 3273  
Db  
2317 CTTTATTAACAAGAAATGACCAACTGTGTATAGCTTGGCTTTCAAGGACATGGACA 2376  
Qy  
3274 AATTTTCAAGTTTGTGTTCTGCCCCTAGACAGTCAATTTCAAAATTTGGGTGTCTTTCTTATG 3333  
Db  
2377 AATTTTCAAGTTTGTGTTCTGCCCCTAGACAGTCAATTTCAAAATTTGGGTGTCTTTCTTATG 2436  
Qy  
3334 GTGTTTCCATGACGACTTTTGAAGAGCTATTTTAAAGCTAGAACTTGAAGCAGAAATTTG 3393  
Db  
2437 GTGTTTCCATGACGACTTTTGAAGAGCTATTTTAAAGCTAGAACTTGAAGCAGAAATTTG 2496  
Qy  
3394 ACCAAGCAGATTTATAGTGTATTTACTCAGAGCCACTGAGGAGAAATTTGATTTCAAAAT 3453  
Db  
2497 ACCAAGCAGATTTATAGTGTATTTACTCAGAGCCACTGAGGAGAAATTTGATTTCAAAAT 2556  
Qy  
3454 CTTTGTGATGAATGAACAGAGCTTACTTTTCTTTCTGAAACCAAGGCTTCTCTAGTGA 3513  
Db  
2557 CTTTGTGATGAATGAACAGAGCTTACTTTTCTTTCTGAAACCAAGGCTTCTCTAGTGA 2616  
Qy  
3514 GCACCATGAGCCTTTTSGAAACCAACAGATGTATACAAATAGCAAGTTTTCATTTCTTTTACT 3573



Db 2617 GCACCATGAGCCTTTGGAAACACACAGATGTATACATAGCAAAAGTTTCATTTCTTACCT 2676  
Qy 3574 TGAACCGTGAAGGTAAATCAGTCAGATCAGTGTGCTTCTGCTTTAAATTTTTCACAG 3633  
Db 2677 TGAACCGTGAAGGTAAATCAGTCAGATCAGTGTGCTTCTGCTTTAAATTTTTCACAG 2736  
Qy 3634 TTCAGATTTTATGTTTTGTTTTCATCAGTCTTTTAAATAATGCTGTGTTTCCCATCAAC 3693  
Db 2737 TTCAGATTTTATGTTTTGTTTTCATCAGTCTTTTAAATAATGCTGTGTTTCCCATCAAC 2796  
Qy 3694 TTGTTCCAGACTTATATTTCTAAACCTGAGACAAACACATATAAATACAAAACAGTC 3753  
Db 2797 TTGTTCCAGACTTATATTTCTAAACCTGAGACAAACACATATAAATACAAAACAGTC 2856  
Qy 3754 TGCTTCTTCAAAATTCGCT-----GACTCAGATATCAGTGAATCTTA 3795  
Db 2857 TGCTTCTTCAAAATTCGCTGGTGAGAGTNNGTGAAGACTCAGATATCAGTGAATCTTA 2916  
Qy 3796 TTAGCTTTTTCACAAGCCAGACATATATGCTGACGATGATTAATGACAGTGAATCTATAT 3855  
Db 2917 TTAGCTTTTTCACAAGCCAGACATATATGCTGACGATGATTAATGACAGTGAATCTATAT 2976  
Qy 3856 CCGTGGCTCCCATAGTGGGCTTTAAATGCTGATGCAATTCAGAAAAGGACTATGTTTTG 3915  
Db 2977 CCGTGGCTCCCATAGTGGGCTTTAAATGCTGATGCAATTCAGAAAAGGACTATGTTTTG 3036  
Qy 3916 CAGCTGTTTTCAACAGTACTATGTTTTTCTTACCTATATAGTGAATATCATTAGTA 3975  
Db 3037 CAGCTGTTTTCAACAGTACTATGTTTTTCTTACCTATATAGTGAATATCATTAGTA 3096  
Qy 3976 ACTACTATCTTTATCATTTAAATGTGACTGAAACCATCCAGATCTGGAGTACCCTATCT 4035  
Db 3097 ACTACTATCTTTATCATTTAAATGTGACTGAAACCATCCAGATCTGGAGTACCCTATCT 3156  
Qy 4036 TTCAGAAATATCTGATATAGTTTTTAAATTTGAGCTGTAATTTTCAAGCAGCTTTGCTTG 4095  
Db 3157 TTCAGAAATATCTGATATAGTTTTTAAATTTGAGCTGTAATTTTCAAGCAGCTTTGCTTG 3216  
Qy 4096 G-AATCATTTGTTACTGCAATGCCACCTTACTTTTGCCATGGAAATGCAGGAATCATAG 4154  
Db 3217 GAAATCATTTGTTACTGCAATGCCACCTTACTCTGCCATGGAAATGCAGGAATCATATA-- 3274  
Qy 4155 ATCAAAGCTTATACTCAACTTAAACTTTTCAAGTCTTTGCACTTCGATATGGATTGGA 4214  
Db 3275 -----AGTCTTTTGCCATCTGCAATTTGGATTGGA 3305  
Qy 4215 CAAGCTGTTGTTGATATCCCTTATTTTATCATTTCTTATTTTGGATGCTAGGAAGCTTA 4274  
Db 3306 CAAGCTGTTGTTGATATCCCTTATTTTATCATTTCTTATTTTGGATGCTAGGAAGCTTA 3365  
Qy 4275 CTGGCATTTTCATTTATGATATTTTATACCTGTAAGTTCCTTCTGCTGTTTTTTCG 4334  
Db 3366 TTGGCATTTTCATTTATGATATATTTTATACCTGTAAGTTCCTTCTGCTGTTTTTTCG 3425  
Qy 4335 CTTATTGTTTATGTTTCCATCAGTTATCTGTTCACTTATATATGCTTTTCACTTTAAG 4394  
Db 3426 CTTATTGTTTATGTTTCCATCAGTTATCTGTTCACTTATATGCTTTTCACTTTAAG 3485  
Qy 4395 AAAATTTTAAATACCAAGAAATTTGGTCAATTTATCTATTTCTGTCGAGCGTTGNCCTGT 4454  
Db 3486 AAAATTTTAAATACCAAGAAATTTGGTCAATTTATCTATCTGTCGAGCGTTGGCTGT 3545  
Qy 4455 ATTGCAATCACTGAATTAATCTTTTATGGGATACAAATTTGCAACTATTTCTTCAATAT 4514  
Db 3546 ATTGCAATCACTGAATTAATCTTTTATGGGATACAAATTTGCAACTATTTCTTCAATAT 3605  
Qy 4515 GCCTTTTGTATCATCATTTCAATCTATCCACTTCTAGTTGCTGATTTCTTTTCAATAAG 4574  
Db 3606 GCCTTTTGTATCATCATTTCAATCTATCCACTTCTAGTTGCTGATTTCTTTTCAATAAG 3665  
Qy 4575 ATTTCTTGGGAAGATGTACGAAATAATGTGCACACCTTATTAATCCATGGGATAGGCTTTCA 4634  
Db 4634 ATTTCTTGGGAAGATGTACGAAATAATGTGCACACCTTATTAATCCATGGGATAGGCTTTCA 4634

Db 3666 ATTTCTTGGGAAGATGTACGAAATAATGTGCACACCTTATTAATCCATGGGATAGGCTTTCA 3725  
Qy 4635 GTAGCTGTTTATATTCGCTTTACCTGCACTGTGTACTGTGATTTTCTCTTTACAAATACTAT 4694  
Db 3726 GTAGCTGTTTATATTCGCTTTACCTGCACTGTGTACTGTGATTTTCTCTTTACAAATACTAT 3785  
Qy 4695 GAGAAAAATATGGAGCGCAGATCAATTAAGAAAAAGATCCCTTTTTCAGAAAACTTTTCAAACG 4754  
Db 3786 GAGAAAAATATGGAGCGCAGATCAATTAAGAAAAAGATCCCTTTTTCAGAAAACTTTTCAAACG 3845  
Qy 4755 AAGTCTTAAAAATAGGAAGCTTCCAGAACCCACAGACAATGAGATGAAGATGTC 4814  
Db 3846 AAGTCTTAAAAATAGGAAGCTTCCAGAACCCACAGACAATGAGATGAAGATGTC 3905  
Qy 4815 AAGTCTGAAAGACTTAAAGGTCAAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAACCA 4874  
Db 3906 AAGTCTGAAAGACTTAAAGGTCAAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAACCA 3965  
Qy 4875 TCCATTATGCTCAGCAATTTGCAATAAAGAAATATGATGACAAGAAAGATTTTCTTTCTTCA 4934  
Db 3966 TCCATTATGCTCAGCAATTTGCAATAAAGAAATATGATGACAAGAAAGATTTTCTTTCTTCA 4025  
Qy 4935 AGAAAAATTAAGAAAGTGGCAACTAAATACATCTCTTCTGCTGTAAGAAAGAGAGATC 4994  
Db 4026 AGAAAAATTAAGAAAGTGGCAACTAAATACATCTCTTCTGCTGTAAGAAAGAGAGATC 4085  
Qy 4995 TTAGGACTATTTGGTCCAAATGGTCTGCAAAAGCACAATTAATTAATTTCTGTTGGT 5054  
Db 4086 TTAGGACTATTTGGTCCAAATGGTCTGCAAAAGCACAATTAATTAATTTCTGTTGGT 4145  
Qy 5055 GATATTGAACCAACTTTCAGGCCAGGTATTTTAGGAGATTTTCTTTCAGACACAAGTGAA 5114  
Db 4146 GATATTGAACCAACTTTCAGGCCAGGTATTTTAGGAGATTTTCTTTCAGACACAAGTGAA 4205  
Qy 5115 GATGATGATTTCACTGAAGTGTATGGTTACTGTCTCAGATAAACCTTTTGTGGCCAGAT 5174  
Db 4206 GATGATGATTTCACTGAAGTGTATGGTTACTGTCTCAGATAAACCTTTTGTGGCCAGAT 4265  
Qy 5175 ACTACATTTCCAGGAACATTTTGAATTTTATGAGAGTGTCAAAGGAATGAGTGAAGTGAC 5234  
Db 4266 ACTACATTTCCAGGAACATTTTGAATTTTATGAGAGTGTCAAAGGAATGAGTGAAGTGAC 4325  
Qy 5235 ATGAAAGAAAGTCATTAAGTCCGAATTAACACATGCACTTGATTTTAAAGAAACATTTTCAGAAG 5294  
Db 4326 ATGAAAGAAAGTCATTAAGTCCGAATTAACACATGCACTTGATTTTAAAGAAACATTTTCAGAAG 4385  
Qy 5295 ACTGTTAAAGAAACTACTCTGAGGAATCAAAACGAAAGTTGTGTTTGTCTTAAGTATGCTA 5354  
Db 4386 ACTGTTAAAGAAACTACTCTGAGGAATCAAAACGAAAGTTGTGTTTGTCTTAAGTATGCTA 4445  
Qy 5355 GGGAAATCCTCAGATTTACTTTGCTAGATGAACCATCTACAGGTATGATCCCAAAGCCAAA 5414  
Db 4446 GGGAAATCCTCAGATTTACTTTGCTAGATGAACCATCTACAGGTATGATCCCAAAGCCAAA 4505  
Qy 5415 CAGCAGATGTGGCGAGCAATTCGAACTGCAATTTTAAAGAAAGCGGCTGCTATTCTG 5474  
Db 4506 CAGCAGATGTGGCGAGCAATTCGAACTGCAATTTTAAAGAAAGCGGCTGCTATTCTG 4565  
Qy 5475 ACCACTCATATATGAGGAGGAGGAGGCTGTCTGTGATCGAGTAGCTATCATGTTGCT 5534  
Db 4566 ACCACTCATATATGAGGAGGAGGAGGCTGTCTGTGATCGAGTAGCTATCATGTTGCT 4625  
Qy 5535 GGGCAGTTAAGATGTATCGGAACAGTACAACTCTAAAGTAAATTTTGGAAAAGGCTAC 5594  
Db 4626 GGGCAGTTAAGATGTATCGGAACAGTACAACTCTAAAGTAAATTTTGGAAAAGGCTAC 4685  
Qy 5595 TTTTTCGAAATTAATTTGAAGGACTGATAGAAAACTAGAAAGTAGACCGCTTTCAAAGA 5654  
Db 4686 TTTTTCGAAATTAATTTGAAGGACTGATAGAAAACTAGAAAGTAGACCGCTTTCAAAGA 4745  
Qy 5655 GAAATTCAGTATATTTTCCCAATGCAAGCCGTCAGGAAAGTTTTTCTTCTATTTTGGCT 5714  
Db 4746 GAAATTCAGTATATTTTCCCAATGCAAGCCGTCAGGAAAGTTTTTCTTCTATTTTGGCT 4805

QY 5715 TATAAAATCCCTAAGGAAGATGTTCTAGTCCCTTTTCAAAATCTTTTAAAGCTGGAAGAA 5774  
 DB 4806 TATAAAATCCCTAAGGAAGATGTTCTAGTCCCTTTTCAAAATCTTTTAAAGCTGGAAGAA 4865  
 QY 5775 GCTAAACATGCTTTTGGCAATGGAAGATATAGCTTTTCTCAAGCAACATTTGGAACAGGTT 5834  
 DB 4866 GCTAAACATGCTTTTGGCAATGGAAGATATAGCTTTTCTCAAGCAACATTTGGAACAGGTT 4925  
 QY 5835 TTTGTAGAACTCACTAAAGAACGAAGAGGAGGAAGATATAGTTGTGGAACTTTAAACAGC 5894  
 DB 4926 TTTGTAGAACTCACTAAAGAACGAAGAGGAGGAAGATATAGTTGTGGAACTTTAAACAGC 4985  
 QY 5895 ACACTTTGGTGGGAACGAACACGAAGAGATAGTAGTATTTTGAATTTGTATTTGGTGG 5954  
 DB 4986 ACACTTTGGTGGGAACGAACACGAAGAGATAGTAGTATTTTGAATTTGTATTTGGTGG 5045  
 QY 5955 TCTGCTTACTGGGACCTCTTTCTTTTCTTTTCACTTAATTTTAACTTTGGTTTAAAGAGTTT 6014  
 DB 5046 TCTGCTTACTGGGACCTCTTTCTTTTCTTTTCACTTAATTTTAACTTTGGTTTAAAGAGTTT 5105  
 QY 6015 TATTGGAATGGTAACCTGGAGAACGAAGAGGACCTTTGAAATTTTCTTAAGCTCCTTAAT 6074  
 DB 5106 TATTGGAATGGTAACCTGGAGAACGAAGAGGACCTTTGAAATTTTCTTAAGCTCCTTAAT 5165  
 QY 6075 GAAATGCTGGTGTGTGTTTCTTTCTTTTAAATTAAGAGTATGTAATTAAGTGA 6134  
 DB 5166 GAAATGCTGGTGTGTGTTTCTTTCTTTTAAATTAAGAGTATGTAATTAAGTGA 5225  
 QY 6135 AGCTGCATGTTTGTATTTGAAGTATATTTGAACATATAGTTTGTATGTCATCTTTTCAAC 6194  
 DB 5226 AGCTGCATGTTTGTATTTGAAGTATATTTGAACATATAGTTTGTATGTCATCTTTTCAAC 5285  
 QY 6195 ATTGAGAAACAGTCTTCTGAATTTGTGATTTAAAGAAATTTGAATAGATAGTTTAT 5254  
 DB 5286 ATTGAGAAACAGTCTTCTGAATTTGTGATTTAAAGAAATTTGAATAGATAGTTTAT 5345  
 QY 6255 TTTAAAGTTATCTTTAAAGTTATGCTATCTTTTAAATTAAGTACGTAAATGTTCCAACTAA 6314  
 DB 5346 TTTAAAGTTATCTTTAAAGTTATGCTATCTTTTAAATTAAGTACGTAAATGTTCCAACTAA 5405  
 QY 6315 ATAAAAAATAATACATACTAATATGCATAGAAAAGATACATAAAGCAATGTGAAAGTT 6372  
 DB 5406 ATAAAAAATAATATCAATACTAATGCATAGAAAAGATACATAAAGCAATGTGAAAGTT 5463

RESULT 4  
 US-09-971-121-5  
 ; Sequence 5, Application US/09971121  
 ; Patent No. US20020111477A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hu, Yi  
 ; APPLICANT: Nepomichy, Boris  
 ; TITLE OF INVENTION: No. US20020111477A1e1 Human Transporter Proteins and Polynucleot  
 ; TITLE OF INVENTION: Same  
 ; FILE REFERENCE: LEX-0250-USA  
 ; CURRENT APPLICATION NUMBER: US/09/971,121  
 ; PRIOR FILING DATE: 2001-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/239,629  
 ; PRIOR FILING DATE: 2000-10-10  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5  
 ; LENGTH: 5262  
 ; TYPE: DNA  
 ; ORGANISM: homo sapiens  
 US-09-971-121-5  
 Query Match 80.43; Score 5247.4; DB 10; Length 5262;  
 Best Local Similarity 99.8; Pred. No. 0;  
 Matches 5248; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
 QY 757 ACTGTTGATATGGTGGTATTTCAAAATCTGCTCTACCTATTTTCAACATGCTTGTACT 816

1 ACTGTTGATATGGTGGTATTTCAAAATCTGCTCTACCTATTTTCAACATGCTTGTACT 60  
 817 TTTTCAAGAGCTGACAGATGCTGCTCCATGCAATCTGCTCCAGTTTCTTAAGAGAGACAGCT 876  
 61 TTTTCAAGAGCTGACAGATGCTGCTCCATGCAATCTGCTCCAGTTTCTTAAGAGAGACAGCT 120  
 877 TGGAGTATGCTTAAATCCATCTTACCTGGGATGAAACACAGCTGCTTATTTTCCGCTTAAAA 936  
 121 TGGAGTATGCTTAAATCCATCTTACCTGGGATGAAACACAGCTGCTTATTTTCCGCTTAAAA 180  
 937 ATTACATGAGTTTATCTGCTGGCTCCGGGTTGTTGTTGTTTCTTCTTCTTAAATAGG 996  
 181 ATTACATGAGTTTATCTGCTGGCTCCGGGTTGTTGTTGTTTCTTCTTCTTAAATAGG 240  
 997 TTTTATTCAGAAAACATGCTCCACTGCAATTTAGGGAGTGGAGTTTGGAGACAGACCAGAA 1056  
 241 TTTTATTCAGAAAACATGCTCCACTGCAATTTAGGGAGTGGAGTTTGGAGACAGACCAGAA 300  
 1057 CACTTCTACTCAAGAAATTTACTTTAAATCAAAATCAGAAACCAAAAGAGTAGTGTTCAGGAAA 1116  
 301 CACTTCTACTGAGAAATTTACTTTAAATCAAAATCAGAAACCAAAAGAGTAGTGTTCAGGAAA 360  
 1117 TTTCTTTTCCACTATTTTCTTTTAAATTTAGGTTAAATTTAATTTAGCATGATGCAATCAATA 1176  
 361 TTTCTTTTCCACTATTTTCTTTTAAATTTAGGTTAAATTTAATTTAGCATGATGCAATCAATA 420  
 1177 AGAATATGAAGAGTGCCTAATATAGAACTCAATCTTATGCAAGAGTTTACTCTTTCTTA 1236  
 421 AGAATATGAAGAGTGCCTAATATAGAACTCAATCTTATGCAAGAGTTTACTCTTTCTTA 480  
 1237 ATCTAATTTCTTGGATATATCTCCAGTGACTAATTTTACAAGCAGCATCATGCAAGAGTGT 1296  
 481 ATCTAATTTCTTGGATATATCTCCAGTGACTAATTTTACAAGCAGCATCATGCAAGAGTGT 540  
 1297 CTACTGATCATCTACTGATGTCATAATTTCTGAAGAAATATACAAATGAAAAGAAATGT 1356  
 541 CTACTGATCATCTACTGATGTCATAATTTCTGAAGAAATATACAAATGAAAAGAAATGT 600  
 1357 TAAATCCAGTCTCTTAAGCCGAGCACTTTTGTAGGTGCTTTTCAAGACGCTCAATGT 1416  
 601 TAAATCCAGTCTCTTAAGCCGAGCACTTTTGTAGGTGCTTTTCAAGACGCTCAATGT 660  
 1417 CCTATGAATCTCGTTTTTCTTCTGATATGATTTCCAGTATCTTCTATTTATATGCAATCAA 1476  
 661 CCTATGAATCTCGTTTTTCTTCTGATATGATTTCCAGTATCTTCTATTTATATGCAATCAA 720  
 1477 GAGTGGCTGTTTCAAAATCATGTGAGGCTGCTCAGTACTGCTCAGGTTTTCAGGTTT 1536  
 721 GAGTGGCTGTTTCAAAATCATGTGAGGCTGCTCAGTACTGCTCAGGTTTTCAGGTTT 780  
 1537 TACAAGCATCCATAGATGCTGCCATTTATACAGTTGAAGCAATGTTTCTCTTTGGAAGG 1596  
 781 TACAAGCATCCATAGATGCTGCCATTTATACAGTTGAAGCAATGTTTCTCTTTGGAAGG 840  
 1597 AGCTGGAGTCAACTAAAGCTGTTTATTTATGGAGAAATCTGCTGTTGAGAAATAGATACCT 1656  
 841 AGCTGGAGTCAACTAAAGCTGTTTATTTATGGAGAAATCTGCTGTTGAGAAATAGATACCT 900  
 1657 TTTCCCGAGGAGTAAATTTAATATACCTAGTTATAGCAATTTTACCTTTTGGATACTTTT 1716  
 901 TTTCCCGAGGAGTAAATTTAATATACCTAGTTATAGCAATTTTACCTTTTGGATACTTTT 960  
 1717 TGGCAATTCATATCGTAGCAGAAAAGAAAATAAAGAAATTTTAAAGATAATAGG 1776  
 961 TGGCAATTCATATCGTAGCAGAAAAGAAAATAAAGAAATTTTAAAGATAATAGG 1020  
 1777 GACTTTCATGATGCTGCTTTTGGCTTTCTCTGGTTTCTTCTATATACAAAGTTTAAATTTTC 1836  
 1021 GACTTTCATGATGCTGCTTTTGGCTTTCTCTGGTTTCTTCTATATACAAAGTTTAAATTTTC 1080  
 1837 TTTATGCTCTTTTATGAGGAGTCAATGCGACAGCTTTCTTTGTTATTTCTCTCAAGTAGCA 1896





QY 4057 TTTTAAATGAGCTGATATTTCAAGCAGCTTTGCTTGAATCATTTGTTACTGCAATGC 4116  
|||||  
Db 3301 TTTTAAATGAGCTGATATTTCAAGCAGCTTTGCTTGAATCATTTGTTACTGCAATGC 3360  
|||||  
QY 4117 CAGCTTACTTTGGCATGGAAAAATGCGAGAAATCATTAAGATCAAAAGCTTATATCTCAACTTA 4176  
|||||  
Db 3361 CAGCTTACTTTGGCATGGAAAAATGCGAGAAATCATTAAGATCAAAAGCTTATATCTCAACTTA 3420  
|||||  
QY 4177 AACCTTTACAGCTCTTTGGCCATCTGCATATTGGAATGCGACAAGCTGTGTTGATATCCCT 4236  
|||||  
Db 3421 AACCTTTACAGCTCTTTGGCCATCTGCATATTGGAATGCGACAAGCTGTGTTGATATCCCT 3480  
|||||  
QY 4237 TATTTTATATCATTTATTTTATGATGCTAGGAAGCTTACTGGCAATTTCAATATGGAATAT 4296  
|||||  
Db 3481 TATTTTATATCATTTATTTTATGATGCTAGGAAGCTTATTTGGCAATTTCAATATGGAATAT 3540  
|||||  
QY 4297 ATTTTATATCTGTAAGTTCCCTGCTGTGGTTTTTTTGGCTTATTTGGTTATGTTCCATCAG 4356  
|||||  
Db 3541 ATTTTATATCTGTAAGTTCCCTGCTGTGGTTTTTTTGGCTTATTTGGTTATGTTCCATCAG 3600  
|||||  
QY 4357 TTATTCGTGTCACCTATATTTGCTTCTTCCACCTTTAAGAAAAATTTTAAATACCAAGAAT 4416  
|||||  
Db 3601 TTATTCGTGTCACCTATATTTGCTTCTTTCACCTTTAAGAAAAATTTTAAATACCAAGAAT 3660  
|||||  
QY 4417 TTTGGTCATTTATCTATTTCTGTGGCAGCGTTGNCCTTGTATTGCAATCACTGAAATAACTT 4476  
|||||  
Db 3661 TTTGGTCATTTATCTATTTCTGTGGCAGCGTTGNCCTTGTATTGCAATCACTGAAATAACTT 3720  
|||||  
QY 4477 TCTTTATGGGATACAAATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCATTTCCAA 4536  
|||||  
Db 3721 TCTTTATGGGATACAAATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCATTTCCAA 3780  
|||||  
QY 4537 TCTATCCACTTTAGTGTGCTGATTTCTTCAATGAATTTCTTGGAGAAATGATACGAA 4596  
|||||  
Db 3781 TCTATCCACTTTAGTGTGCTGATTTCTTCAATGAATTTCTTGGAGAAATGATACGAA 3840  
|||||  
QY 4597 AAAATGFGGACACTATAATCCATGGGATAGGCTTTTCAGTAGCTGTATTATCGCCTTACC 4656  
|||||  
Db 3841 AAAATGFGGACACTATAATCCATGGGATAGGCTTTTCAGTAGCTGTATTATCGCCTTACC 3900  
|||||  
QY 4657 TGCAGTGTGACTGTGGATTTTCTCTTACAATACATATGAGAAAAAATATGAGGCGAGAT 4716  
|||||  
Db 3901 TGCAGTGTGACTGTGGATTTTCTCTTACAATACATATGAGAAAAAATATGAGGCGAGAT 3960  
|||||  
QY 4717 CAATAAGAAAGATCCCTTTTTCAGAAACCTTTTCAAGAGTCTAAATATGGAAGCTTC 4776  
|||||  
Db 3961 CAATAAGAAAGATCCCTTTTTCAGAAACCTTTTCAAGAGTCTAAATATGGAAGCTTC 4020  
|||||  
QY 4777 CAGAACCCAGACAATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGACTAAAGGTCA 4836  
|||||  
Db 4021 CAGAACCCAGACAATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGACTAAAGGTCA 4080  
|||||  
QY 4837 AAGAGCTGATGGGTGCGCAGTGTGTGAGGAGAAACCATCATATATGGTCAGCAATTTGC 4896  
|||||  
Db 4081 AAGAGCTGATGGGTGCGCAGTGTGTGAGGAGAAACCATCATATATGGTCAGCAATTTGC 4140  
|||||  
QY 4897 ATAAAGAAATATGACAAAGAAATTTTCTTTCAGAAAGATTAAGAAAGTGGCAA 4956  
|||||  
Db 4141 ATAAAGAAATATGACAAAGAAATTTTCTTTCAGAAAGATTAAGAAAGTGGCAA 4200  
|||||  
QY 4957 CTAATAACATCTCTTTCTGTGTGAAAAAGGAGAGATCTTTAGGACTATTGGGTCCAAATG 5016  
|||||  
Db 4201 CTAATAACATCTCTTTCTGTGTGAAAAAGGAGAGATCTTTAGGACTATTGGGTCCAAATG 4260  
|||||  
QY 5017 GTGCTGCGCAAAAGCACAAATTTAATAATTTCTGGTTGGTGATATGAAACCAACTTCAGGCC 5076  
|||||  
Db 4261 GTGCTGCGCAAAAGCACAAATTTAATAATTTCTGGTTGGTGATATGAAACCAACTTCAGGCC 4320  
|||||  
QY 5077 AGGTATTTTATAGGATATTTCTTCAGAGACAAGTGAAGATGATTTCACTGAAGTGA 5136  
|||||  
Db 4321 AGGTATTTTATAGGATATTTCTTCAGAGACAAGTGAAGATGATTTCACTGAAGTGA 4380  
|||||  
QY 5137 TGGGTTACTGTCCTCAGATAAAACCCCTTTGTGGCCAGATACATATTGCAGGAACATTTTG 5196  
|||||

Db 4381 TGGGTTACTGTCCTCAGATAAAACCCCTTTGTGGCCAGATACATATTGCAGGAACATTTTG 4440  
|||||  
QY 5197 AAAATTTATGGAGCTGTCAAAGGAATGAGTGAAGTGCATCAAGAAAGCTCATTAAGTCGAA 5256  
|||||  
Db 4441 AAAATTTATGGAGCTGTCAAAGGAATGAGTGAAGTGCATCAAGAAAGCTCATTAAGTCGAA 4500  
|||||  
QY 5257 TAAACATGCGCACTTGATTTTAAAAAGAAACAATCTTTCAAGAACATCTGTTAAAGAAACCTACCTGCGAG 5316  
|||||  
Db 4501 TAAACATGCGCACTTGATTTTAAAAAGAAACAATCTTTCAAGAACATCTGTTAAAGAAACCTACCTGCGAG 4560  
|||||  
QY 5317 GAATCAAAACGAAAGTGTGTTTTGCTCTAAGTATGCTTAGGGAATCCTCAGATTTACTTTGC 5376  
|||||  
Db 4561 GAATCAAAACGAAAGTGTGTTTTGCTCTAAGTATGCTTAGGGAATCCTCAGATTTACTTTGC 4620  
|||||  
QY 5377 TAGATGAACCATCTACAGGTATGGATCCCAAAGCCAAACAGCACATGTGGGAGCAATTC 5436  
|||||  
Db 4621 TAGATGAACCATCTACAGGTATGGATCCCAAAGCCAAACAGCACATGTGGGAGCAATTC 4680  
|||||  
QY 5437 GAACTGCAATTTAAAAACAGAAAGCGGCTGCTATTCTGACCACCTCACTATATGAGGAGG 5496  
|||||  
Db 4681 GAACTGCAATTTAAAAACAGAAAGCGGCTGCTATTCTGACCACCTCACTATATGAGGAGG 4740  
|||||  
QY 5497 CAGAGGCTGTCTGTGAATCGAGTAGCTATCATGTGTCTGCGGAGTTAAGATGTATCGGAA 5556  
|||||  
Db 4741 CAGAGGCTGTCTGTGATCGAGTAGCTATCATGTGTCTGCGGAGTTAAGATGTATCGGAA 4800  
|||||  
QY 5557 CAGTACACATCTAAGAGTAAATTTGGAAAAGGCTACTTTTGGAAAATTTAAATTTGAAGG 5616  
|||||  
Db 4801 CAGTACACATCTAAGAGTAAATTTGGAAAAGGCTACTTTTGGAAAATTTAAATTTGAAGG 4860  
|||||  
QY 5617 ACTGATAGAAAACCTAGAACTAGACCCGCTTCAAAGAGAAATTCAGTATATTTTCCCAA 5676  
|||||  
Db 4861 ACTGATAGAAAACCTAGAACTAGACCCGCTTCAAAGAGAAATTCAGTATATTTTCCCAA 4920  
|||||  
QY 5677 ATGCAAGCCGTGAGAAAAGTTTCTCTATTTTGGCTTTATAAAATTCCTTAAGGAAGATG 5736  
|||||  
Db 4921 ATGCAAGCCGTGAGAAAAGTTTCTCTATTTTGGCTTTATAAAATTCCTTAAGGAAGATG 4980  
|||||  
QY 5737 TTCAGTCCCTTTCACATCTTTTAAAGCTGGAGAGAGCTAAACATGCTTTTGCCTTTC 5796  
|||||  
Db 4981 TTCAGTCCCTTTCACATCTTTTAAAGCTGGAGAGAGCTAAACATGCTTTTGCCTTTC 5040  
|||||  
QY 5797 AAGATATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAGAAC 5856  
|||||  
Db 5041 AAGATATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAGAAC 5100  
|||||  
QY 5857 AAGAGGAGAAATATAGTTGTGGAACTTTTAAACAGCACACTTTGGTGGGAAACGAAACAC 5916  
|||||  
Db 5101 AAGAGGAGAAATATAGTTGTGGAACTTTTAAACAGCACACTTTGGTGGGAAACGAAACAC 5160  
|||||  
QY 5917 AAGAGATAGAGTAGTATTTGAAATTTGATTTGCTGCTGCTTACCTGGGACTTCTTTC 5976  
|||||  
Db 5161 AAGAGATAGAGTAGTATTTGAAATTTGATTTGTTGGTCTGCTTACCTGGGACTTCTTTC 5220  
|||||  
QY 5977 TTTTTCACCTAATTTTAACTTTGTTGTTAAAAAGTTTTT 6014  
|||||  
Db 5221 TTTTTCACCTAATTTTAACTTTGTTTAAAAAGTTTTT 5258  
|||||

RESULT 5

US-09-971-121-1  
; Sequence 1, Application US/09971121  
; Patent No. US20020111477A1  
; GENERAL INFORMATION:  
; APPLICANT: Hu, Yi  
; APPLICANT: Nepomichy, Boris  
; TITLE OF INVENTION: No. US20020111477A1el Human Transporter Proteins and Polynucleot  
; TITLE OF INVENTION: Same  
; FILE REFERENCE: LEX-0250-USA  
; CURRENT APPLICATION NUMBER: US/09/971,121  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/239,629

;; PRIOR FILING DATE: 2000-10-10  
;; NUMBER OF SEQ ID NOS: 5  
;; SOFTWARE: Fast-Seq for Windows Version 4.0  
;; SEQ ID NO 1  
;; LENGTH: 4929  
;; TYPE: DNA  
;; ORGANISM: homo sapiens  
US-09-971-121-1

Query Match 75.4%; Score 4918.4; DB 10; Length 4929;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 4919; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1011 ATGTCCACTGCAATAGGAGGAGTAGGAGTTGGAGACAGACCAGAACACTTCTACTGAAG 1070  
Db 1 ATGTCCACTGCAATAGGAGGAGTAGGAGTTGGAGACAGACCAGAACACTTCTACTGAAG 60

Qy 1071 AATTACTTAATTAATCGAGAACCAAAAAGAGTAGTGTTCAGGAATCTTTTTCCACTA 1130  
Db 61 AATTACTTAATTAATCGAGAACCAAAAAGAGTAGTGTTCAGGAATCTTTTTCCACTA 120

Qy 1131 TTTTCTTTTATTTGGTTAAATTAATTAATAGCATGATGCATCCAAATAGAAATATGAGAA 1190  
Db 121 TTTTCTTTTATTTGGTTAAATTAATTAATAGCATGATGCATCCAAATAGAAATATGAGAA 180

Qy 1191 GTGCTAATATAGAACTCAATCTCTATGCAAGTTTACTCTTTCTAATCTAAATCTTTGGA 1250  
Db 181 GTGCTAATATAGAACTCAATCTCTATGCAAGTTTACTCTTTCTAATCTAAATCTTTGGA 240

Qy 1251 TATATCTCCAGTGACTAATATTAACAAGCAGCATCATGAGAAAGTGCTACTGATCATCTA 1310  
Db 241 TATATCTCCAGTGACTAATATTAACAAGCAGCATCATGAGAAAGTGCTACTGATCATCTA 300

Qy 1311 CCTGATGTCATTAATTAAGAGATATACAAATGAAAGAAATGTTAAATCCAGTCTC 1370  
Db 301 CCTGATGTCATTAATTAAGAGATATACAAATGAAAGAAATGTTAAATCCAGTCTC 360

Qy 1371 TCTAAGCCGAGCAACTTTGTAGGTGTGTTTCAAGACTCCATGCTCTATGAATCTCGT 1430  
Db 361 TCTAAGCCGAGCAACTTTGTAGGTGTGTTTCAAGACTCCATGCTCTATGAATCTCGT 420

Qy 1431 TTTTCTTCTGATATGATTCAGATATCTCTATTAATATGATTAACAAGCTGGCTGTCA 1490  
Db 421 TTTTCTTCTGATATGATTCAGATATCTCTATTAATATGATTAACAAGCTGGCTGTCA 480

Qy 1491 AAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGTTTCACAGTTTACAGCATCCATA 1550  
Db 481 AAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGTTTTCACAGTTTACAGCATCCATA 540

Qy 1551 GATGCTGCCATTATACAGTTTGAAGACCAATGTTTCTCTTTGGAAGAGCTGGAGTCAACT 1610  
Db 541 GATGCTGCCATTATACAGTTTGAAGACCAATGTTTCTCTTTGGAAGAGCTGGAGTCAACT 600

Qy 1611 AAAGCTGTTATATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTA 1670  
Db 601 AAAGCTGTTATATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTA 660

Qy 1671 ATTTTAATATACCTAGTATAGCAATTTTACCTTTTGATACCTTTTGGCAATTCATATC 1730  
Db 661 ATTTTAATATACCTAGTATAGCAATTTTACCTTTTGATACCTTTTGGCAATTCATATC 720

Qy 1731 GTACAGAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACT 1790  
Db 721 GTACAGAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACT 780

Qy 1791 GCCTTTTGGCTTTCCTCGGTTCTCTATATACAAAGTTTAAATTTTCTATGTCCTCTCT 1850  
Db 781 GCCTTTTGGCTTTCCTCGGTTCTCTATATACAAAGTTTAAATTTTCTATGTCCTCTCT 840

Qy 1851 ATGCGAGTCATTTGGACAGCTTCTTTGTATTTTCTCCTCAAGTAGCAGCATTTGATATTT 1910  
Db 841 ATGCGAGTCATTTGGACAGCTTCTTTGTATTTTCTCCTCAAGTAGCAGCATTTGATATTT 900

1911 CTGCTTTTTTCTTTATGGAATATCACTGTAATTTTTTGTCTTTTAAGTCAACCTCTTT 1970  
901 CTGCTTTTTTCTTTATGGAATATCACTGTAATTTTTTGTCTTTTAAGTCAACCTCTTT 960

Qy 1971 TTTAAAAATCAAAACATGTGGGAATAGTTGAATTTTTTGTACTGTGGCTTTTGGATTT 2030  
Db 961 TTTAAAAATCAAAACATGTGGGAATAGTTGAATTTTTTGTACTGTGGCTTTTGGATTT 1020

Qy 2031 ATTGGCTTTATGATATCTCTATAGAAAGTTTTCCTCAATCGTTAGTGTGGCTTTTTCAGT 2090  
Db 1021 ATTGGCTTTATGATATCTCTATAGAAAGTTTTCCTCAATCGTTAGTGTGGCTTTTTCAGT 1080

Qy 2091 CCTTCTGTCACTGTACTTTTGTGATTTGTCACAGCTCATGATTTAGAAATTTT 2150  
Db 1081 CCTTCTGTCACTGTACTTTTGTGATTTGTCACAGCTCATGATTTAGAAATTTT 1140

Qy 2151 AATGAAGGTGCTTCATTTTCAAAATTTGACTGCGAGGCCATATCTCTAATTTATTAATTT 2210  
Db 1141 AATGAAGGTGCTTCATTTTCAAAATTTGACTGCGAGGCCATATCTCTAATTTATTAATTT 1200

Qy 2211 ATCATGCTCACACTTAATAGTATATCTATGCTCTCTGTGCTCTATCTTGTATCAAGTC 2270  
Db 1201 ATCATGCTCACACTTAATAGTATATCTATGCTCTCTGTGCTCTATCTTGTATCAAGTC 1260

Qy 2271 ATTCCAGGGGAAATTTGGCTTTACGGAGATCATCTTTATATTTTCTGAAGCCTTCATATTGG 2330  
Db 1261 ATTCCAGGGGAAATTTGGCTTTACGGAGATCATCTTTATATTTTCTGAAGCCTTCATATTGG 1320

Qy 2331 TCAAGAGTAAAGAAATTTATGAGGAGTTATCAGAGGGCAATGTTTAATGGAATATTTAGT 2390  
Db 1321 TCAAGAGTAAAGAAATTTATGAGGAGTTATCAGAGGGCAATGTTTAATGGAATATTTAGT 1380

Qy 2391 TTTTAGTGAATATTTTTCAGCCAGTTTCTTCAGAAATTTGTAGGAAAGAGCCATTAAGAAAT 2450  
Db 1381 TTTTAGTGAATATTTTTCAGCCAGTTTCTTCAGAAATTTGTAGGAAAGAGCCATTAAGAAAT 1440

Qy 2451 AGTGGTATTTCAGAAAGACATACAGAAAGAGGGTGAATAATGTTGAGGCTTTGAGAAATTTG 2510  
Db 1441 AGTGGTATTTCAGAAAGACATACAGAAAGAGGGTGAATAATGTTGAGGCTTTGAGAAATTTG 1500

Qy 2511 TCATTTGACATATATGAGGGTCAAGATTACTGCTTACTTTGGCCACAGTGGAAACAGGAAAG 2570  
Db 1501 TCATTTGACATATATGAGGGTCAAGATTACTGCTTACTTTGGCCACAGTGGAAACAGGAAAG 1560

Qy 2571 AGTACATTTGATGAATATTTCTTGTGACTCTGCCCACCTTCTCATGGTTTGCATCTATA 2630  
Db 1561 AGTACATTTGATGAATATTTCTTGTGACTCTGCCCACCTTCTCATGGTTTGCATCTATA 1620

Qy 2631 TATGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAAAATGATTGGCAAT 2690  
Db 1621 TATGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAAAATGATTGGCAAT 1680

Qy 2691 TGTCCACAGTTAGATATACCTTTGATGTTTTCAGAGTAGAAGAAATTTTATCAATTTTG 2750  
Db 1681 TGTCCACAGTTAGATATACCTTTGATGTTTTCAGAGTAGAAGAAATTTTATCAATTTTG 1740

Qy 2751 GCTTCAATCAAGGGATACAGCCCAATATATCAAGAGTGCAGAGGTTTTTACTA 2810  
Db 1741 GCTTCAATCAAGGGATACAGCCCAATATATCAAGAGTGCAGAGGTTTTTACTA 1800

Qy 2811 GATTTAGACATGACAGTACTCAAGATATCAAGCTTAAATAATTAAGTGGTGGTTCAAAAA 2870  
Db 1801 GATTTAGACATGACAGTACTCAAGATATCAAGCTTAAATAATTAAGTGGTGGTTCAAAAA 1860

Qy 2871 AGAAAGCTGCTATTAGGAATGCTGTTTGGGAAACCCAAAGATATCTGCTGTAGATGAA 2930  
Db 1861 AGAAAGCTGCTATTAGGAATGCTGTTTGGGAAACCCAAAGATATCTGCTGTAGATGAA 1920

Qy 2931 CCAACAGCTGGAATGACCCCTGTTCTCGACATATTTGTATGTAATCTTTTAAATATCAGA 2990  
Db 1921 CCAACAGCTGGAATGACCCCTGTTCTCGACATATTTGTATGTAATCTTTTAAATATCAGA 1980

Qy 2991 AAAGCCAAATCGGCTGACAGTGTTCAGTACTCTATTTTATGATGAAGTCAATTTCTTGCA 3050

1981 AAAGCCAATCGGTGACAGTGTTCAGTACTCAITTCATGGATGAAGCTGACATTCITGCA 2040  
3051 GATAGGAAAGCTGTGATATACAAAGGAATGCTGAAATGTGTGGTTCCTCAATGTCCCTC 3110  
2041 GATAGGAAAGCTGTGATATACAAAGGAATGCTGAAATGTGTGGTTCCTCAATGTCCCTC 2100  
3111 AAAAGTAAATGGGGATCGGCTACCGCTCAGCATGTACATACAGAAATATGTGCGACA 3170  
2101 AAAAGTAAATGGGGATCGGCTACCGCTCAGCATGTACATACAGAAATATGTGCGACA 2160  
3171 GAATCTCTTTCTTCACTGGTTAAACACATATACCTGGAGCTACTTTATTAACAAGAAAT 3230  
2161 GAATCTCTTTCTTCACTGGTTAAACACATATACCTGGAGCTACTTTATTAACAAGAAAT 2220  
3231 GACCAACAACCTGTGTATAGCTTGCCTTCAAGGACATGAGCAAAATTTTCAGGTGTGTTT 3290  
2221 GACCAACAACCTGTGTATAGCTTGCCTTCAAGGACATGAGCAAAATTTTCAGGTGTGTTT 2280  
3291 TCTGCCCTAGACAGTCAATCAAAATTTGGGTGTCATTTCTTATGGTGTTCCTCATGACGACT 3350  
2281 TCTGCCCTAGACAGTCAATCAAAATTTGGGTGTCATTTCTTATGGGTGTTCCTCATGACGACT 2340  
3351 TTGGAAGACGCTATTTTAAAGCTAGAGTTGAAGCAGAAATTGACCAAGCAGATTAAGT 3410  
2341 TTGGAAGACGCTATTTTAAAGCTAGAGTTGAAGCAGAAATTGACCAAGCAGATTAAGT 2400  
3411 GTATTTACTCAGCAGCCACTGGAGGAAGAAATGGATCAAAATCTTTTGTGAAATGGAA 3470  
2401 GTATTTACTCAGCAGCCACTGGAGGAAGAAATGGATCAAAATCTTTTGTGAAATGGAA 2460  
3471 CAGAGCTTACTTCTTCTGAAACCAAGGCTCTCTAGTGAGCACCATGAGCCCTTGG 3530  
2461 CAGAGCTTACTTCTTCTGAAACCAAGGCTCTCTAGTGAGCACCATGAGCCCTTGG 2520  
3531 AAACAACAGATGTATACAAATAGCAAGTTTCATTTCTTTACCTTGAAACCTGAAAGTAAA 3590  
2521 AAACAACAGATGTATACAAATAGCAAGTTTCATTTCTTTACCTTGAAACCTGAAAGTAAA 2580  
3591 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTAAATTTTCAAGTCTGCTTCTTCAAAATTCCT 3650  
2581 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTAAATTTTCAAGTCTGAGATTTTATGTTT 2640  
3651 TTGCTTCACTCCTTTTAAATGCTGTGGTTCCTCATCAAACTGTTCAGACTTATAT 3710  
2641 TTGCTTCACTCCTTTTAAATGCTGTGGTTCCTCATCAAACTGTTCAGACTTATAT 2700  
3711 TTTCTAAACCTGGAGACAAACACATATAAATAACAAACCAAGTCTGCTTCTTCAAAATTCCT 3770  
2701 TTTCTAAACCTGGAGACAAACACATATAAATAACAAACCAAGTCTGCTTCTTCAAAATTCCT 2760  
3771 GCTGACTCAGATATCAGTGATCTTATAGCTTTTCAAGGCGAGCAATATGAGTGGTGG 3830  
2761 GCTGACTCAGATATCAGTGATCTTATAGCTTTTCAAGGCGAGCAATATGAGTGGTGG 2820  
3831 ATGATTAATCAGAGTACTATGATCGTGGCTCCCATAGTGGGCTTTAAATGTGATG 3890  
2821 ATGATTAATCAGAGTACTATGATCGTGGCTCCCATAGTGGGCTTTAAATGTGATG 2880  
3891 CAITTCAGAAAAGGACTATGTTTTTGACAGCTGTTTTCAACAGTACTATGTTTTTCTTTA 3950  
2881 CAITTCAGAAAAGGACTATGTTTTTGACAGCTGTTTTCAACAGTACTATGTTTTTCTTTA 2940  
3951 CCTATATAGTAATATCATATAGTACTATCTTTTATCAATTAATATGACTGAAC 4010  
2941 CCTATATAGTAATATCATATAGTACTATCTTTTATCAATTAATATGACTGAAC 3000  
4011 ATCCAGATCTGGAGTACCCATCTTTTCAAGAAATTAAGTACTATGTTTTTAAATTTGAG 4070  
3001 ATCCAGATCTGGAGTACCCATCTTTTCAAGAAATTAAGTACTATGTTTTTAAATTTGAG 3060  
4071 CTGTATTTTCAAGCAGCTTGTGCTTGAATCATTTGTTACTGCAATGCCCTTACTTTGCT 4130

3061 CTGTATTTTCAAGCAGCTTTGCTTTGGAATCATTTGTTACTGCAATGCCACTTTACTTTGCC 3120  
4131 ATGGAATATGAGAGATCAATAAGATCAAGCTTATATCTCAACTTTAAACTTTTCAAGTCTT 4190  
3121 ATGGAATATGAGAGATCAATAAGATCAAGCTTATATCTCAACTTTAAACTTTTCAAGTCTT 3180  
4191 TTGCCATCTGCATATTTGGAATTTGGCAAGCTGTGTGATATCCCTTATTTTATCATTT 4250  
3181 TTGCCATCTGCATATTTGGAATTTGGCAAGCTGTGTGATATCCCTTATTTTATCATTT 3240  
4251 CTTATTTTGTATGCTAGAGAGCTTACTGGCATTTCAATATGGAATATATATTTTATATCTGA 4310  
3241 CTTATTTTGTATGCTAGAGAGCTTACTGGCATTTCAATATGGAATATATATTTTATATCTGA 3300  
4311 AAGTTCTCTGCTGTGCTTTTTCCTTATTTGCTTATGTTTCCATCAGTTATTTCTGTTCACT 4370  
3301 AAGTTCTCTGCTGTGCTTTTTCCTTATTTGCTTATGTTTCCATCAGTTATTTCTGTTCACT 3360  
4371 TATATTTGCTTCTTTCACCTTTTAAAGAAATTTTAAATACCAAGAAATTTTGTTCATTTATC 4430  
3361 TATATTTGCTTCTTTCACCTTTTAAAGAAATTTTAAATACCAAGAAATTTTGTTCATTTATC 3420  
4431 TATTCCTGCGAGCGCTTGNCTTGTATGCAATCACTGAAATAAATTTCTTTTATGGAATAC 4490  
3421 TATTCCTGCGAGCGCTTGGCTTGTATGCAATCACTGAAATAAATTTCTTTTATGGAATAC 3480  
4491 ACAATTCGAATCTATTTCTTCAATTTGCTTGTATGTTTCCATCATTCCCAATCTTCTTA 4550  
3481 ACAATTCGAATCTATTTCTTCAATTTGCTTGTATGTTTCCATCATTCCCAATCTTCTTA 3540  
4551 GGTTCCTCTGATTTCTTTCATAAAGATTTCTTGGGAAGATTTGACGAAAAATTTGTGACACC 4610  
3541 GGTTCCTCTGATTTCTTTCATAAAGATTTCTTGGGAAGATTTGACGAAAAATTTGTGACACC 3600  
4611 TATAATTCATGGAATAGGCTTTCAGTAGCTGTATATCGCTTACCTGAGTGTGACTG 4670  
3601 TATAATTCATGGAATAGGCTTTCAGTAGCTGTATATCGCTTACCTGAGTGTGACTG 3660  
4671 TGGATTTTCTCTTACATCTATGAGAAAAATATGAGGAGCAATCAATTAAGAAAAAGAT 4730  
3661 TGGATTTTCTCTTACATCTATGAGAAAAATATGAGGAGCAATCAATTAAGAAAAAGAT 3720  
4731 CCCTTTTTCAGAAAACCTTTTCAACGAAGTCTAAAAATAGGAAGCTTCCAGAACCCAGAC 4790  
3721 CCCTTTTTCAGAAAACCTTTTCAACGAAGTCTAAAAATAGGAAGCTTCCAGAACCCAGAC 3780  
4791 AATGAGATGAAGATGAAGATGTCAAGCTGAAAGACTAAAGTCAAAAGCTGATGGGT 4850  
3781 AATGAGATGAAGATGAAGATGTCAAGCTGAAAGACTAAAGTCAAAAGCTGATGGGT 3840  
4851 TGCCAGTCTTGTGAGGAGAAACCATCCATTTATGTTGAGCAATTTGTCATAAAGATATGAT 4910  
3841 TGCCAGTCTTGTGAGGAGAAACCATCCATTTATGTTGAGCAATTTGTCATAAAGATATGAT 3900  
4911 GACAAAAAGATTTTCTTTTCAAGAAAAAGTAAAGAAAGTGGCAACTTAAATACATCTCT 4970  
3901 GACAAAAAGATTTTCTTTTCAAGAAAAAGTAAAGAAAGTGGCAACTTAAATACATCTCT 3960  
4971 TTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTTGGGTCAAAATGCTGGGCAAAAGC 5030  
3961 TTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTTGGGTCAAAATGCTGGGCAAAAGC 4020  
5031 ACAATTTATATTTCTGTTGGTGTATTTGAACTTCAACCACTTCAGGCCAGGTATTTTATGGA 5090  
4021 ACAATTTATATTTCTGTTGGTGTATTTGAACTTCAACCACTTCAGGCCAGGTATTTTATGGA 4080  
5091 GATTTATCTTCAGAGACAAGTGAAGATGATTCATCTGAAAGTGTATGGGTCTGCTCT 5150  
4081 GATTTATCTTCAGAGACAAGTGAAGATGATTCATCTGAAAGTGTATGGGTCTGCTCT 4140  
5151 CAGATAAACCCCTTTTGTGCCAGATACACTCATTTGAGGAAATTTTGAATTTATGAGCT 5210  
4141 CAGATAAACCCCTTTTGTGCCAGATACACTCATTTGAGGAAATTTTGAATTTATGAGCT 4200







Db 3121 AATGCCACCCCTACTCTCCCATGGAAATGACAGAAATCATATA----- 3161  
Qy 4172 ACTTAAACTTTTCAGGCTCTTTTGGCATCTGCATATTTGGAATGGACAAGCTGTGTGTATAT 4231  
Db 3162 -----AGGCTTTTGGCATCTGCATATTTGGAATGGACAAGCTGTGTGTATAT 3209  
Qy 4232 CCCCTTATTTTATCATCTTATTTTGTATGCTAGGAAGCTTACTGCGCATTTCAATTATGG 4291  
Db 3210 CCCCTTATTTTATCATCTTATTTTGTATGCTAGGAAGCTTATTTGCGCATTTCAATTATGG 3269  
Qy 4292 ATTATATTTTATCTGTAAGTTTCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTC 4351  
Db 3270 ATTATATTTTATCTGTAAGTTTCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTC 3329  
Qy 4352 ATCAGTTTATTTCTGTCTATATATGCTTCTTTCACCTTTTAAAGAAATTTTAAATACCAA 4411  
Db 3330 ATCAGTTTATTTCTGTCTATATATGCTTCTTTCACCTTTTAAAGAAATTTTAAATACCAA 3389  
Qy 4412 AGAATTTTGTCTATTTATCTTCTGTGGCAGGTTGCTTGTATTTGCAATCACTGAAT 4471  
Db 3390 AGAATTTTGTCTATTTATCTTCTGTGGCAGGTTGCTTGTATTTGCAATCACTGAAT 3449  
Qy 4472 AACTTTCTTTATGGATACACAATTTCAATTTCTTCTATTAAGAAATTTTGGAAAGATGT 4531  
Db 3450 AACTTTCTTTATGGATACACAATTTCAATTTCTTCTATTAAGAAATTTTGGAAAGATGT 3509  
Qy 4532 TCCAATCTATCCACTTCTAGGTTGCTGTATTTCTTCTATTAAGAAATTTTGGAAAGATGT 4591  
Db 3510 TCCAATCTATCCACTTCTAGGTTGCTGTATTTCTTCTATTAAGAAATTTTGGAAAGATGT 3569  
Qy 4592 ACGAAAAATTTGGACACCTTATTAATCCATGGATAGGCTTTTCACTGTATATCGCC 4651  
Db 3570 ACGAAAAATTTGGACACCTTATTAATCCATGGATAGGCTTTTCACTGTATATCGCC 3629  
Qy 4652 TTACCTGCTAGTGTACTGTGGATTTTCTTCTTACATATCTATGAGAAAAATATGGAG 4711  
Db 3630 TTACCTGCTAGTGTACTGTGGATTTTCTTCTTACATATCTATGAGAAAAATATGGAG 3689  
Qy 4712 CAGATCAATAAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAAGAAAGTCTAAAAATAGGAA 4771  
Db 3690 CAGATCAATAAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAAGAAAGTCTAAAAATAGGAA 3749  
Qy 4772 GCTTCCAGAACCCAGACCAATGAGATGAAGATGAAGATGTCAAAGCTCAAAGACTTAA 4831  
Db 3750 GCTTCCAGAACCCAGACCAATGAGATGAAGATGAAGATGTCAAAGCTCAAAGACTTAA 3809  
Qy 4832 GGTCAAAGAGCTGATGGTTGCCAGTGTGTGAGGAGAAACCATCTATATGTGACGAA 4891  
Db 3810 GGTCAAAGAGCTGATGGTTGCCAGTGTGTGAGGAGAAACCATCTATATGTGACGAA 3869  
Qy 4892 TTTGCATAAAGAAATATGATGACAAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGT 4951  
Db 3870 TTTGCATAAAGAAATATGATGACAAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGT 3929  
Qy 4952 GGCAACTTAAATACATCTCTTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTTGGGTCC 5011  
Db 3930 GGCAACTTAAATACATCTCTTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTTGGGTCC 3989  
Qy 5012 AAATGGTCTGGCAAAAGCAATTTATATTTCTGGTTGGTGATATTTGAACCACTTC 5071  
Db 3990 AAATGGTCTGGCAAAAGCAATTTATTAATTTCTGGTTGGTGATATTTGAACCACTTC 4049  
Qy 5072 AGGCGAGTATTTTATGAGATATTTCTTCCAGAGACAAGTGAAGATGATGATTTCACTGAA 5131  
Db 4050 AGGCGAGTATTTTATGAGATATTTCTTCCAGAGACAAGTGAAGATGATGATTTCACTGAA 4109  
Qy 5132 GTGTATGGGTACTGTCTCAGATAAACCCCTTTTGTGGCCAGATCTACTATTTGAGGAACA 5191  
Db 4110 GTGTATGGGTACTGTCTCAGATAAACCCCTTTTGTGGCCAGATCTACTATTTGAGGAACA 4169  
Qy 5192 TTTTGAATTTATGGAGCTGTCAAAGAAATGAGTGAAGTGAATGAAGAAAGTCAATAG 5251  
Db 4170 TTTTGAATTTATGGAGCTGTCAAAGAAATGAGTGAAGTGAATGAAGAAAGTCAATAG 4229

Qy 5252 TCGAATAACACATGCACTTTGATTTTAAAGAACATCTTTCAGAGACTGTAAAGAACTACC 5311  
Db 4230 TCGAATAACACATGCACTTTGATTTTAAAGAACATCTTTCAGAGACTGTAAAGAACTACC 4289  
Qy 5312 TCGAGGAATCAAAACGAAAGTTGTGTGTCTTAAGTATGCTAGGAAATCTCAGATTAC 5371  
Db 4290 TCGAGGAATCAAAACGAAAGTTGTGTGTCTTAAGTATGCTAGGAAATCTCAGATTAC 4349  
Qy 5372 TTTGCTAGATGAACATCTACAGGTATGATCCCAAGCCAAACACACATGTGGCGAGC 5431  
Db 4350 TTTGCTAGATGAACATCTACAGGTATGATCCCAAGCCAAACACACATGTGGCGAGC 4409  
Qy 5432 AATTCGAACCTGCAATTTTAAAAACAGAAAGCGGCTGCTATTCTGACCACTCATATATGA 5491  
Db 4410 AATTCGAACCTGCAATTTTAAAAACAGAAAGCGGCTGCTATTCTGACCACTCATATATGA 4469  
Qy 5492 GAGGCGAGAGGCTGTCTGTGATCGAGTAGCTATCATGTTGCTGCGGAGTTAAGATGTAT 5551  
Db 4470 GAGGCGAGAGGCTGTCTGTGATCGAGTAGCTATCATGTTGCTGCGGAGTTAAGATGTAT 4529  
Qy 5552 CGGAACAGTACAACTTAAAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTAAT 5611  
Db 4530 CGGAACAGTACAACTTAAAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTAAT 4589  
Qy 5612 GAAGGACTGGATAGAAAACTTAGAGTAGACCGCTTCAAAGAGAAATTCAGTATATTT 5671  
Db 4590 GAAGGACTGGATAGAAAACTTAGAGTAGACCGCTTCAAAGAGAAATTCAGTATATTT 4649  
Qy 5672 CCCAAATGAAGCGTCAAGAAAGTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGA 5731  
Db 4650 CCCAAATGAAGCGTCAAGAAAGTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGA 4709  
Qy 5732 AGATGTTTCACTCCCTTTCACATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGC 5791  
Db 4710 AGATGTTTCACTCCCTTTCACATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGC 4769  
Qy 5792 CATTGAAGAAATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAA 5851  
Db 4770 CATTGAAGAAATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAA 4829  
Qy 5852 AGAACAGAGGAGGAAGATTAATAGTTGTGAACTTTTAAACAGCACACTTTGTGGGAACG 5911  
Db 4830 AGAACAGAGGAGGAAGATTAATAGTTGTGAACTTTTAAACAGCACACTTTGTGGGAACG 4889  
Qy 5912 AACACAAGAGATAGTAGTATTTTGA 5939  
Db 4890 AACACAAGAGATAGTAGTATTTTGA 4917

## RESULT 7

US-09-971-121-3  
; Sequence 3, Application US/09971121  
; Patent No. US20020111477A1  
; GENERAL INFORMATION:  
; APPLICANT: Hu, Yi  
; APPLICANT: Nepomnichy, Boris  
; TITLE OF INVENTION: Same  
; TITLE OF INVENTION: Same  
; FILE REFERENCE: LEX-0250-USA  
; CURRENT APPLICATION NUMBER: US/09/971,121  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/239,629  
; PRIOR FILING DATE: 2000-10-10  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 4785  
; TYPE: DNA  
; ORGANISM: homo sapiens

US-09-971-121-3

Query Match 72.9%; Score 4754.8; DB 10; Length 4785;

Best Local Similarity 99.8%; Pred. No. 0; Matches 4756; Conservative 4; Mismatches 7; Indels 0; Gaps 0;											
QY	1011	ATGTC	CACTG	CAATTA	GGAGGTAGG	AGTTGG	AGACAG	ACCAG	ACCACTTCTACTGAAG	1070	
DB	1	ATGTC	CACTG	CAATTA	GGAGGTAGG	AGTTGG	AGACAG	ACCAG	ACCACTTCTACTGAAG	60	
QY	1071	AAATTA	CTTAATTA	AAATTC	CAAGAAC	CAAAAG	AGTAGT	GTTCAG	GAATAATCTTTTCCACTA	1130	
DB	61	AAATTA	CTTAATTA	AAATTC	CAAGAAC	CAAAAG	AGTAGT	GTTCAG	GAATAATCTTTTCCACTA	120	
QY	1131	TTTTTT	TTTATTTT	GGTTAA	TATTAAT	TAGCAT	GATG	CCATCA	AAATAGAAATATGAAGAA	1190	
DB	121	TTTTTT	TTTATTTT	GGTTAA	TATTAAT	TAGCAT	GATG	CCATCA	AAATAGAAATATGAAGAA	180	
QY	1191	GTGCT	TAATATAG	AACTCA	ATCTT	TATG	GACAAG	TTTACT	CTTTCTAATCTAATCTTTGGA	1250	
DB	181	GTGCT	TAATATAG	AACTCA	ATCTT	TATG	GACAAG	TTTACT	CTTTCTAATCTAATCTTTGGA	240	
QY	1251	TATACT	CCAGTG	ACTAAT	TATTA	CAAG	CAGCAT	CATG	CAGAAAGTGTCTACTGATCATCTA	1310	
DB	241	TATACT	CCAGTG	ACTAAT	TATTA	CAAG	CAGCAT	CATG	CAGAAAGTGTCTACTGATCATCTA	300	
QY	1311	CCTGAT	GTCA	TAATTA	CTGA	AGAT	TATAC	AAATG	AAAGAAATGTTTAAACATCCAGTCTC	1370	
DB	301	CCTGAT	GTCA	TAATTA	CTGA	AGAT	TATAC	AAATG	AAAGAAATGTTTAAACATCCAGTCTC	360	
QY	1371	TCTAAG	CCGAG	CAACTTT	TAGT	GTGGT	TTTCA	AGAC	TCCATGCTTATGAACTTCGT	1430	
DB	361	TCTAAG	CCGAG	CAACTTT	TAGT	GTGGT	TTTCA	AGAC	TCCATGCTTATGAACTTCGT	420	
QY	1431	TTTTTT	CTCAT	TATG	ATTC	CAGT	ATCT	CTAT	TATATGATTTCAAGAGCTGGCTGTCA	1490	
DB	421	TTTTTT	CTCAT	TATG	ATTC	CAGT	ATCT	CTAT	TATATGATTTCAAGAGCTGGCTGTCA	480	
QY	1491	AAAT	CATG	TGAG	GTG	CTC	CAGT	TTTCA	AGTTTTACAAGCATCCATA	1550	
DB	481	AAAT	CATG	TGAG	GTG	CTC	CAGT	TTTCA	AGTTTTACAAGCATCCATA	540	
QY	1551	GATG	CTGCC	AT	TACAG	TGTTT	TTTCA	AGAC	CTTGGAGAGCTGGAGTCAACT	1610	
DB	541	GATG	CTGCC	AT	TACAG	TGTTT	TTTCA	AGAC	CTTGGAGAGCTGGAGTCAACT	600	
QY	1611	AAAG	CTGTT	AT	AGCG	AAACT	GCTCT	TG	AGAAATAGATACATCTTTCCCGCAGGAGTA	1670	
DB	601	AAAG	CTGTT	AT	AGCG	AAACT	GCTCT	TG	AGAAATAGATACATCTTTCCCGCAGGAGTA	660	
QY	1671	ATTTT	AAATAT	ACCTAG	TTATAG	CAATTTT	CCACT	TTTGG	AAATTTTCAATATC	1730	
DB	661	ATTTT	AAATAT	ACCTAG	TTATAG	CAATTTT	CCACT	TTTGG	AAATTTTCAATATC	720	
QY	1731	GTAG	CAGAAA	AGAAAA	AAAAAT	TTTAA	AGAT	TAAT	GGGACTTCAATGATCT	1790	
DB	721	GTAG	CAGAAA	AGAAAA	AAAAAT	TTTAA	AGAT	TAAT	GGGACTTCAATGATCT	780	
QY	1791	GCCTTT	TGGCTTT	CTCGG	TTCTCT	TATAC	AGTTT	TAAT	TTTCTTATGCTCCCTCTT	1850	
DB	781	GCCTTT	TGGCTTT	CTCGG	TTCTCT	TATAC	AGTTT	TAAT	TTTCTTATGCTCCCTCTT	840	
QY	1851	ATGG	CAGT	CAAT	TTGG	CAGCTTT	TTGTT	TATTT	CTCTCAAGTAGCAGCAATTTGATATTT	1910	
DB	841	ATGG	CAGT	CAAT	TTGG	CAGCTTT	TTGTT	TATTT	CTCTCAAGTAGCAGCAATTTGATATTT	900	
QY	1911	CTGCT	TTTTTT	CTTTT	ATG	ATAT	ATCAT	CTGT	ATTTT	TTTGGCTTTAATCTGACACCTCTT	1970
DB	901	CTGCT	TTTTTT	CTTTT	ATG	ATAT	ATCAT	CTGT	ATTTT	TTTGGCTTTAATCTGACACCTCTT	960
QY	1971	TTTT	AAAA	AAAT	CAAA	CATG	TGG	GAATAG	TTTCTTCTGCTTTTGGATTT	2030	
DB	961	TTTT	AAAA	AAAT	CAAA	CATG	TGG	GAATAG	TTTCTTCTGCTTTTGGATTT	1020	
QY	2031	ATTG	GCCT	TTAT	GAAT	TAAT	CTCT	CAT	AGAAAGTTTTTCCAAATCGTTAGTGTGGCTTTTCACT	2090	





Db	4321	TTGTGTTTCTCTAAGTATGCTAGGAATCCCTCAGATTACTTTTGTCTAGATGAACCATCT	4380
Qy	5391	ACAGGTATGATCCCAAGCAACCAAGCAGCAGATGTCGCGAGCAATTCGAATGCATTTAA	5450
Db	4381	ACAGGTATGATCCCAAGCAACCAAGCAGCAGATGTCGCGAGCAATTCGAATGCATTTAA	4440
Qy	5451	ACAGAAAGCGGCTGCTATCTTCTGACCACTACCTATATGAGGAGGAGCAGAGGCTGCTGT	5510
Db	4441	ACAGAAAGCGGCTGCTATCTTCTGACCACTACCTATATGAGGAGGAGCAGAGGCTGCTGT	4500
Qy	5511	GATCGAGTATGATCATGCTGCTGCGCAGTAAAGATGATCGGAACAGTACAACTATA	5570
Db	4501	GATCGAGTATGATCATGCTGCTGCGCAGTAAAGATGATCGGAACAGTACAACTATA	4560
Qy	5571	ACAGTAATTTGGAAGGCTACTTTTGGAAATTTAAATGGAAGGCTGGAATAGAAAC	5630
Db	4561	ACAGTAATTTGGAAGGCTACTTTTGGAAATTTAAATGGAAGGCTGGAATAGAAAC	4620
Qy	5631	CTAGAGTACACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCAAGCCGTCAG	5690
Db	4621	CTAGAGTACACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCAAGCCGTCAG	4680
Qy	5691	GAAGATTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTGAGTCCCTTTCA	5750
Db	4681	GAAGATTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTGAGTCCCTTTCA	4740
Qy	5751	CAATCTTTTAAAGCTGGAAGGCT 5777	
Db	4741	CAATCTTTTAAAGCTGGAAGGCT 4767	

RESULT 8

US-10-108-260A-160  
 ; Sequence 160, Application US/10108260A  
 ; Publication No. US20040005560A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HELIX RESEARCH INSTITUTE  
 ; TITLE OF INVENTION: No. US20040005560A1e1 full length cdna  
 ; FILE REFERENCE: HI-A0106  
 ; - CURRENT APPLICATION NUMBER: US/10/108,260A  
 ; CURRENT FILING DATE: 2002-03-27  
 ; NUMBER OF SEQ ID NOS: 5458  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 160  
 ; LENGTH: 3347  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-108-260A-160

Query Match	49.98;	Score 3258.2;	DB 12;	Length 3347;
Best Local Similarity	98.54;	Pred. No. 0;		
Matches 3327;	Conservative 0;	Mismatches 4;	Indels 47;	Gaps 2;
Qy	2999	TCGGGTGACAGTGTTCAGTACTCATTTTCAGTGAAGCTGACATCTTTGCGATAGGAA	3058	
Db	1	TCGGGTGACAGTGTTCAGTACTCATTTTCAGTGAAGCTGACATCTTTGCGATAGGAA	60	
Qy	3059	AGCTGTGATATCAAGGAATGCTGAATGTTGTTGTTTCAATGTTCTCAAAAGTAA	3118	
Db	61	AGCTGTGATATCAAGGAATGCTGAATGTTGTTGTTTCAATGTTCTCAAAAGTAA	120	
Qy	3119	ATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGGCCACAGAACTCTCT	3178	
Db	121	ATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGGCCACAGAACTCTCT	180	
Qy	3179	TTCTTCACTGGTTAAACAAATATACCTGGAGCTACTTTTATTAACAGAAATGACCAACA	3238	
Db	181	TTCTTCACTGGTTAAACAAATATACCTGGAGCTACTTTTATTAACAGAAATGACCAACA	240	
Qy	3239	ACTTGTATAGCTTGCCTTTCAAGGACATGGAACAAATTTTCAGGTTTGTTCGCGCT	3298	
Db	241	ACTTGTATAGCTTGCCTTTCAAGGACATGGAACAAATTTTCAGGTTTGTTCGCGCT	300	

Qy	3299	AGACAGTCATTCAAATTTGGGTGTCATTTCTTATGGTGTTCATGACGACTTTTGAAGA	3358
Db	301	AGACAGTCATTCAAATTTGGGTGTCATTTCTTATGGTGTTCATGACGACTTTTGAAGA	360
Qy	3359	CGTATTTTAAAGCTAGAGTTGAAGCGAGAAATTCACCAAGCAGATTTAGTGTATTAC	3418
Db	361	CGTATTTTAAAGCTAGAGTTGAAGCGAGAAATTCACCAAGCAGATTTAGTGTATTAC	420
Qy	3419	TCAGCAGCCACTGGAGGAGAAATGGAATCAAAATCTTTTGTATGAAATGGAAACAGACTT	3478
Db	421	TCAGCAGCCACTGGAGGAGAAATGGAATCAAAATCTTTTGTATGAAATGGAAACAGACTT	480
Qy	3479	ACTTATTTCTTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTGGAACAACA	3538
Db	481	ACTTATTTCTTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTGGAACAACA	540
Qy	3539	GATGTATACATAGCAAAAGTTTCATTTCTTTTACCTTGAAACGTGAAAGTAATCAGTGAG	3598
Db	541	GATGTATACATAGCAAAAGTTTCATTTCTTTTACCTTGAAACGTGAAAGTAATCAGTGAG	600
Qy	3599	ATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTTCGTTCA	3658
Db	601	ATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTTCGTTCA	660
Qy	3659	TCACCTCTTTTAAATATGCTGTGGTTTCCCATCAAACTTCTCCAGACTTATATTTCTAAA	3718
Db	661	TCACCTCTTTTAAATATGCTGTGGTTTCCCATCAAACTTCTCCAGACTTATATTTCTAAA	720
Qy	3719	ACCTGGAGACAAACCAATAAATACAAACAAAGTCTGCTTCTTCAAAATTCGTCT	3773
Db	721	ACCTGGAGACAAACCAATAAATACAAACAAAGTCTGCTTCTTCAAAATTCGTCTGCTGA	780
Qy	3774	-----GACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCCAGACATAA	3822
Db	781	GAGTGTGGAAGACTCAGATATCTGTGATCTTATTAGCTTTTTCACAGCCAGACATAA	840
Qy	3823	TGGTGACGATGATTAATGACAGTGTACTATGTATCGTGGCTCCCATAGTGGCGTTTAA	3882
Db	841	TGGTGACGATGATTAATGACAGTGTACTATGTATCGTGGCTCCCATAGTGGCGTTTAA	900
Qy	3883	ATGTGATGATTCAGAAAGGACTATGTTTTCAGCTGTTTTCAGAGTACTATGGTTT	3942
Db	901	ATGTGATGATTCAGAAAGGACTATGTTTTCAGCTGTTTTCAGAGTACTATGGTTT	960
Qy	3943	ATTTCTTTACCTATATTTAGTGAATATCATTAGTAACTACTATCTTTTCAATTTAAATGTA	4002
Db	961	ATTTCTTTACCTATATTTAGTGAATATCATTAGTAACTACTATCTTTTCAATTTAAATGTA	1020
Qy	4003	CTGAAACCATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAATCTGATATAGTTTAA	4062
Db	1021	CTGAAACCATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAATCTGATATAGTTTAA	1080
Qy	4063	AAATTTAGCTGTATTTTCAAGCAGCTTTGCTGGAAATCATTTGTTTAACTGCAATGCCACTT	4122
Db	1081	AAATTTAGCTGTATTTTCAAGCAGCTTTGCTGGAAATCATTTGTTTAACTGCAATGCCACTT	1140
Qy	4123	ACTTTGCCATGGAATTCGAGAGATCATAGATCAAGCTTATATCTCACTTAACTTAACTTT	4182
Db	1141	ACTTTGCCATGGAATTCGAGAGATCATAGATCAAGCTTATATCTCACTTAACTTAACTTT	1170
Qy	4183	CAGGTCTTTTGGCATCTGCATATTCGATTTGGAAGCTGTTTGTGATATCCCTTATTTT	4242
Db	1171	-AGGTCTTTTGGCATCTGCATATTCGATTTGGAAGCTGTTTGTGATATCCCTTATTTT	1229
Qy	4243	TTATCATTTTATTTTGTAGTGTAGAACTTCTGCGATTTTCAATTTGATATATTTTTT	4302
Db	1230	TTATCATTTTATTTTGTAGTGTAGAACTTCTGCGATTTTCAATTTGATATATTTTTT	1289
Qy	4303	ATACTGTAAGTTCCTTCTGCTGTTTTCGCTTATTTGCTTATTTGCTTATCTCAGTATATTC	4362
Db	1290	ATACTGTAAGTTCCTTCTGCTGTTTTCGCTTATTTGCTTATTTGCTTATCTCAGTATATTC	1349



; APPLICANT: SATO, HIROYUKI		; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA	
; APPLICANT: ISHII, SHIZUKO		; FILE REFERENCE: 084335/0160	
; APPLICANT: YAMAMOTO, JUN-ICHI		; CURRENT APPLICATION NUMBER: US/10/094,749	
; APPLICANT: ISONO, YUUKO		; CURRENT FILING DATE: 2002-03-12	
; APPLICANT: HIO, YURI		; PRIOR APPLICATION NUMBER: 60/350,435	
; APPLICANT: OTSUKA, KAORU		; PRIOR FILING DATE: 2002-01-24	
; APPLICANT: NAGAI, KEIICHI		; PRIOR APPLICATION NUMBER: JP 2001-328381	
; APPLICANT: IRIE, RYOTARO		; PRIOR FILING DATE: 2001-09-14	
; APPLICANT: TAMECHIKA, ICHIRO		; NUMBER OF SEQ ID NOS: 3381	
; APPLICANT: SEKI, NAOHIKO		; SOFTWARE: Patencin Ver. 2.1	
; APPLICANT: YOSHIKAWA, TSUTOMU		; SEQ ID NO 984	
; APPLICANT: OTSUKA, MOTOUKI		; LENGTH: 3268	
; APPLICANT: NAGAHARI, KENJI		; TYPE: DNA	
; APPLICANT: MASUHO, YASUHIKO		; ORGANISM: Homo sapiens	
; FILE OF INVENTION: NOVEL FULL-LENGTH cDNA		US-10-094-749-984	
Query Match 48.7%; Score 3178; DB 13; Length 3268;			
Best Local Similarity 98.8%; Pred. No. 0;			
Matches 3224; Conservative 0; Mismatches 6; Indels 33; Gaps 1;			
QY	3102	ATGTTCTCTCAAAAGTAAATGGGGATCGGTACCGCTGAGCATGTACATAGACAAATAT	3161
DB	6	ATGTTCTCTCAAAAGTAAATGGGGATCGGTACCGCTGAGCATGTACATAGACAAATAT	65
QY	3162	TGTGCCACAGAACTCTTTCTTCACTGGTTAAACACATATACCTGGAGCTACTTTATTA	3221
DB	66	TGTGCCACAGAACTCTTTCTTCACTGGTTAAACACATATACCTGGAGCTACTTTATTA	125
QY	3222	CAACAGAAATGACCAAACTTTGTGTATAGTCTGCCCTTCAAGGACATGGAACAAATTTCA	3281
DB	126	CAACAGAAATGACCAAACTTTGTGTATAGTCTGCCCTTCAAGGACATGGAACAAATTTCA	185
QY	3282	G-----GTTTGTCTTCCCTTAGACAGTCAT	3308
DB	186	GGAATTGCTGTATAGACAAAGGATGTTATGTATGTTTCTGCCCTTAGACAGTCAT	245
QY	3309	TCAAAATTTGGGTGTCATTTCTTATGTTTCCATGACGATTTGGAAGACGTTATTTTA	3368
DB	246	TCAAAATTTGGGTGTCATTTCTTATGTTTCCATGACGATTTGGAAGACGTTATTTTA	305
QY	3369	AAGCTAGAAGTTGAAGCAGAAATTTGACCAAGCAGATTATAGTGTATTTACTCAGCAGCCA	3428
DB	306	AAGCTAGAAGTTGAAGCAGAAATTTGACCAAGCAGATTATAGTGTATTTACTCAGCAGCCA	365
QY	3429	CTGGAGGAGAAATGGATTCAAAATCTTTTGATGAAATGGAACAGAGCTTACTTTCTT	3488
DB	366	CTGGAGGAGAAATGGATTCAAAATCTTTTGATGAAATGGAACAGAGCTTACTTTCTT	425
QY	3489	TCTGAAACCAAGGCTTCTAGTGAGCACCATGAGCTTTGGAAACAAACAGATGTATACA	3548
DB	426	TCTGAAACCAAGGCTTCTAGTGAGCACCATGAGCTTTGGAAACAAACAGATGTATACA	485
QY	3549	ATAGCAAAAGTTTCATTTCTTTACCTTGAAACGTTGAAAGTAAATCAGTGAGATCAGTGTG	3608
DB	486	ATAGCAAAAGTTTCATTTCTTTACCTTGAAACGTTGAAAGTAAATCAGTGAGATCAGTGTG	545
QY	3609	CTTCTGCTTTTAAATTTTTTTCACAGTTTCAGATTTTTTATGTTTGGTTCATCTTTT	3668
DB	546	CTTCTGCTTTTAAATTTTTTTCACAGTTTCAGATTTTTTATGTTTGGTTCATCTTTT	605

QY	3669	AAAAATGCTGTGTTGCCATCAAAACCTTGTTCCAGACTTATATTTTCTAAAACTGGAGAC	3722
DB	606	AAAAATGCTGTGTTGCCATCAAAACCTTGTTCCAGACTTATATTTTCTAAAACTGGAGAC	665
QY	3729	AAACCACATAAATACAAAAACAAGTCTGCTCTTTCAAAATTCCTGCTGACTCAGATATCACT	3788
DB	666	AAACCACATAAATACAAAAACAAGTCTGCTCTTTCAAAATTCCTGCTGACTCAGATATCACT	725
QY	3789	GATCTTATTAGCTTTTTTTCACAAGCCAGAAACATAAATGGTGACGATGATTAATCACAAGTGAC	3848
DB	726	GATCTTATTAGCTTTTTTTCACAAGCCAGAAACATAAATGGTGACGATGATTAATCACAAGTGAC	785
QY	3849	TATGATATCCGTGGTCCCATAGTGGGGCTTTAAATGTGATGATTCAGAAAAAGACTAT	3908
DB	786	TATGATATCCGTGGTCCCATAGTGGGGCTTTAAATGTGATGATTCAGAAAAAGACTAT	845
QY	3909	GTTTTTGACGCTGTTTTTCAACAGTACTATGTTTTTATTTCTTTTACCCTATATTTAGTGAATATC	3968
DB	846	GTTTTTGACGCTGTTTTTCAACAGTACTATGTTTTTATTTCTTTTACCCTATATTTAGTGAATATC	905
QY	3969	ATTAGTAACTACTATCTTTTATCAATTTTAAATGTGACTGAAACCATCCAGATCTGGAGTACC	4028
DB	906	ATTAGTAACTACTATCTTTTATCAATTTTAAATGTGACTGAAACCATCCAGATCTGGAGTACC	965
QY	4029	CCATTTCTTCAAGAAATTTACTGATATAGTTTTTAAATTTGAGCTGTATTTTCAACAGCT	4088
DB	966	CCATTTCTTCAAGAAATTTACTGATATAGTTTTTAAATTTGAGCTGTATTTTCAACAGCT	1025
QY	4089	TTGCTTGGAAATCAATGTTACTGCAATGCCACCTTACTTTGCCATGGAAATCCAGAGAAAT	4148
DB	1026	TTGCTTGGAAATCAATGTTACTGCAATGCCACCTTACTTTGCCATGGAAATCCAGAGAAAT	1085
QY	4149	CATAAGATCAAAAGCTTATACTCAACTTAAACTTTTCAAGTCTTTTGGCACTCTGCATATTGG	4208
DB	1086	CATAAGATCAAAAGCTTATACTCAACTTTAACTTTTCAAGTCTTTTGGCACTCTGCATATTGG	1145
QY	4209	ATTGGACAAGCTGTGTTGATATCCCTTTATTTTTTATCAATTTTATTTTGTGATGCTTAGGA	4268
DB	1146	ATTGGACAAGCTGTGTTGATATCCCTTTATTTTTTATCAATTTTATTTTGTGATGCTTAGGA	1205
QY	4269	AGCTTACTGGCAATTCATTTATGGAATTTATTTTTTATCACTGTAAGTTCCTGCTGTGGTT	4328
DB	1206	AGCTTACTGGCAATTCATTTATGGAATTTATTTTTTATCACTGTAAGTTCCTGCTGTGGTT	1265
QY	4329	TTTTGGCTTATGTTTATGTTCCATCAGTTATTTCTGTTCACTTATATGCTTCTTTTCCAC	4388
DB	1266	TTTTGGCTTATGTTTATGTTCCATCAGTTATTTCTGTTCACTTATATGCTTCTTTTCCAC	1325
QY	4389	TTTTAAGAAAAATTTTAAATACCAAAGAAATTTTGGTCAATTTATCTATTCTGTGGCAGCGTTG	4448
DB	1326	TTTTAAGAAAAATTTTAAATACCAAAGAAATTTTGGTCAATTTATCTATTCTGTGGCAGCGTTG	1385
QY	4449	NCTTGTATTGCAATCACTGAAATTAACCTTTCTTTATGGGATACAAATTCGCAACTATTCTT	4508
DB	1386	NCCTTGTATTGCAATCACTGAAATTAACCTTTCTTTATGGGATACAAATTCGCAACTATTCTT	1445
QY	4509	CATTATGCTTTTGTATCATCATTTCCAAATCTATCTAGTTTCCCTGATTTCTTTCTTC	4568
DB	1446	CATTATGCTTTTGTATCATCATTTCCAAATCTATCTAGTTTCCCTGATTTCTTTCTTC	1505
QY	4569	ATAAAGATTTCTTTGGAGAAATGTACGAAAAAATGTGGACACCTTATTAATCCATGGGATAGG	4628
DB	1506	ATAAAGATTTCTTTGGAGAAATGTACGAAAAAATGTGGACACCTTATTAATCCATGGGATAGG	1565
QY	4629	CTTTTCACTAGTGTATATTCGCTTACCTGCACTGTGTACTGTGGATTTTCTCTTTACAA	4688
DB	1566	CTTTTCACTAGTGTATATTCGCTTACCTGCACTGTGTACTGTGGATTTTCTCTTTACAA	1625
QY	4689	TACTATCAGAAAAAATATGGAGCAGATCAATTAAGAAAAAGATCCCTTTTTCAGAAACCTT	4748
DB	1626	TACTATCAGAAAAAATATGGAGCAGATCAATTAAGAAAAAGATCCCTTTTTCAGAAACCTT	1685

QY	3669	AAAAATGCTGTGGTTCCCATCAAACTTGTTCAGACTTATATTTTCTAAAACTGGAGAC	3728
DB	606	AAAAATGCTGTGGTTCCCATCAAACTTGTTCAGACTTATATTTTCTAAAACTGGAGAC	665
QY	3729	AAACCACTAAATACAAAACAAAGTCTGCTTCTTCAAAATTCGTGCTGACTCAGATATCAGT	3788
DB	666	AAACCACTAAATACAAAACAAAGTCTGCTTCTTCAAAATTCGTGCTGACTCAGATATCAGT	725
QY	3789	GATCTTATTAGCTTTTTCACAAGCCAGAACATAATGGTGACGATGATTAAATGACAGTAC	3848
DB	726	GATCTTATTAGCTTTTTCACAAGCCAGAACATAATGGTGACGATGATTAAATGACAGTAC	785
QY	3849	TATGTATCCGTGGCTCCCATAGTGGCGCTTTTAAATGTGATGCTATTTCAGAAAAGSACTAT	3908
DB	786	TATGTATCCGTGGCTCCCATAGTGGCGCTTTTAAATGTGATGCTATTTCAGAAAAGSACTAT	845
QY	3909	GTTTTTGCAGCTGTTTTTCAACAGTACTATAGTGTATTTTCTTTACCTATATATAGTAATATC	3968
DB	846	GTTTTTGCAGCTGTTTTTCAACAGTACTATAGTGTATTTTCTTTACCTATATATAGTAATATC	905
QY	3969	ATTAGTAACTACTATCTTTTATCAATTTAAATGTGACTGAAACCATCCAGATCTGGAGTACC	4028
DB	906	ATTAGTAACTACTATCTTTTATCAATTTAAATGTGACTGAAACCATCCAGATCTGGAGTACC	965
QY	4029	CCATTTCTTTCAAGAAATTTACTGTATATAGTTTAAAAATTTGAGCTGTATTTTCAAGCAGCT	4088
DB	966	CCATTTCTTTCAAGAAATTTACTGTATATAGTTTAAAAATTTGAGCTGTATTTTCAAGCAGCT	1025
QY	4089	TTGCTTGGAAATCATTTGTATCTGCAATGCCACTTACTTTTGCCTGCAATGCGAGAGAT	4148
DB	1026	TTGCTTGGAAATCATTTGTACTGCAATGCCACTTACTTTTGCCTGCAATGCGAGAGAT	1085
QY	4149	CATAAGATCAAGCTTATACCTCAACTTAACTTTTAACTTTTGCCTGCAATGCGATTTGG	1145
DB	1086	CATAAGATCAAGCTTATACCTCAACTTAACTTTTAACTTTTGCCTGCAATGCGATTTGG	1145
QY	4209	ATTGGAACAAGCTGTTTGTATATCCCTTATTTTATCAATTTTATTTTATGATGCTAGGA	4268
DB	1146	ATTGGAACAAGCTGTTTGTATATCCCTTATTTTATCAATTTTATTTTATGATGCTAGGA	1205
QY	4269	AGCTTACTGCAATTTCAATATGGAATATATTTTATCTGCTGCTGCTGCTGCTGCTGCTGCT	4328
DB	1206	AGCTTACTGCAATTTCAATATGGAATATATTTTATCTGCTGCTGCTGCTGCTGCTGCTGCT	1265
QY	4329	TTTTGCTTATTTGTTATGTTTCCATCAGTTATCTGTTTCACTTATATGTTTCTTTTACC	4388
DB	1266	TTTTGCTTATTTGTTATGTTTCCATCAGTTATCTGTTTCACTTATATGTTTCTTTTACC	1325
QY	4389	TTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATCTATTTCTGTCGACGCTTG	4448
DB	1326	TTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATCTATTTCTGTCGACGCTTG	1385
QY	4449	NCCTGTATTGCAATCAGTGAATAAATCTTTTATGGAATACAAATGCAACTATTTCTT	4508
DB	1386	GCTTGTATTGCAATCAGTGAATAAATCTTTTATGGAATACAAATGCAACTATTTCTT	1445
QY	4509	CATTATGCTTTTGTATCATCTTCCATCTTCACTTCTAGGTTGCTGATTTCTTTTC	4568
DB	1446	CATTATGCTTTTGTATCATCTTCCATCTTCACTTCTAGGTTGCTGATTTCTTTTC	1505
QY	4569	ATAAAGATTCTTGGGAAGAAATGTACGAAATAATGTGACACCTATAATCCATGGGATAGG	4628
DB	1506	ATAAAGATTCTTGGGAAGAAATGTACGAAATAATGTGACACCTATAATCCATGGGATAGG	1565
QY	4629	CTTTTCACTAGCTGTTTATATCGCTTACCTGAGTGTGTACTGTGGAATTTCTCTTTCAAA	4688
DB	1566	CTTTTCACTAGCTGTTTATATCGCTTACCTGAGTGTGTACTGTGGAATTTCTCTTTCAAA	1625
QY	4689	TACTATGAGAAATAATATGAGGAGAGATCAATTAAGAAAGATCCCTTTTTCAGAAACCTT	4748
DB	1626	TACTATGAGAAATAATATGAGGAGAGATCAATTAAGAAAGATCCCTTTTTCAGAAACCTT	1685



Query Match	27.4%;	Score 1789.4;	DB 11;	Length 3928;
Best Local Similarity	98.2%;	Pred. No. 0;		
Matches 1833;	Conservative	0;	Mismatches	1;
			Indels	33;
			Gaps	1;
QY	1941	GTATTTTTCCTTTTAATGCTGCACACCTCTCTTTTAAAAAATCAAAACATGTGGGAATAGTT	2000	
DB	1	GTATTTTTCCTTTTAATGCTGCACACCTCTCTTTTAAAAAATCAAAACATGTGGGAATAGTT	60	
QY	2001	GAATTTTTCCTTACTGTGGCTTTTGGATTTATTGGCCCTTAATGATAATATCCCTCATAGAAAAGT	2060	
DB	61	GAATTTTTCCTTACTGTGGCTTTTGGATTTATTGGCCCTTATGATAATATCCCTCATAGAAAAGT	120	
QY	2061	TTTTCGCAAAATCGTTAGTGTGGCTTTTCAGTGCCCTTTCTGTCACTGTACTTTGTGTGATTTGGT	2120	
DB	121	TTTTCGCAAAATCGTTAGTGTGGCTTTTCAGTGCCCTTTCTGTCACTGTACTTTGTGTGATTTGGT	180	
QY	2121	ATTGCACAGGTCATGCATTTTAGAAGATTTTAAATGAAGGTGCTTCATTTTCAAAATTTTGACT	2180	
DB	181	ATTGCACAGGTCATGCATTTTAGAAGATTTTAAATGAAGGTGCTTCATTTTCAAAATTTTGACT	240	
QY	2181	GCAGGCCCATATCCTCTAAATTATTACAATTTATCATGTCTCACCTTAATAGTATATTCTAT	2240	
DB	241	GCAGGCCCATATCCTCTAAATTATTACAATTTATCATGTCTCACCTTAATAGTATATTCTAT	300	
QY	2241	GTCTCTTTGGCTGTCTATCTTTGATCAAGTCATTTCCAGGGAAATTTGGCTTACGGAGATCA	2300	
DB	301	GTCTCTTTGGCTGTCTATCTTTGATCAAGTCATTTCCAGGGAAATTTGGCTTACGGAGATCA	360	
QY	2301	TCCTTATATTTTCTGAAGCCTTCATATTTGGTCAAAGAGTAAAGAAATATTATGAGGAGTTA	2360	
DB	361	TCCTTATATTTTCTGAAGCCTTCATATTTGGTCAAAGAGTAAAGAAATATTATGAGGAGTTA	420	
QY	2361	TCAGAGGCAATGTTAATGCAAAATATTAGTTTTAGTCAAAATATTATGAGCCAGATTTCTTCA	2420	
DB	421	TCAGAGGCAATGTTAATGCAAAATATTAGTTTTAGTCAAAATATTATGAGCCAGATTTCTTCA	480	
QY	2421	GAATTTTGTAGGAAAAGAGCCATAAAGAAATTTAGTGGTATTTCAGAGACATACAGAAAAGG	2480	
DB	481	GAATTTTGTAGGAAAAGAGCCATAAAGAAATTTAGTGGTATTTCAGAGACATACAGAAAAGG	540	
QY	2481	GGTGAAAATGTGAGGCTTTTGAGAAATTTGTCAATTTGCATATATATGAGGCTCAGATTACT	2540	
DB	541	GGTGAAAATGTGAGGCTTTTGAGAAATTTGTCAATTTGCATATATATGAGGCTCAGATTACT	600	
QY	2541	GCCTTACTTGGCCACACGTGGAAACAGGAAAAGATGATTTGATGAAATATCTTTTGTGGACTC	2600	
DB	601	GCCTTACTTGGCCACACGTGGAAACAGGAAAAGATGATTTGATGAAATATCTTTTGTGGACTC	660	
QY	2601	TGCCACCTTCTGTANTGGTTTGCATCTATATATGGAACACAGAGTCTCAGAAAATAGATGAA	2660	
DB	661	TGCCACCTTCTGTANTGGTTTGCATCTATATATGGAACACAGAGTCTCAGAAAATAGATGAA	720	
QY	2661	ATGTTTCAAGCAGAAAATGATTTGGCATTTGTCCACAGTTTAGATATATACACTTTTGATGTT	2720	
DB	721	ATGTTTCAAGCAGAAAATGATTTGGCATTTGTCCACAGTTTAGATATATACACTTTTGATGTT	780	
QY	2721	TTGACAGTAGAAGAAAATTTATCAATTTTGGCTTCAATCAAGGGATACACAGCCACAAT	2780	
DB	781	TTGACAGTAGAAGAAAATTTATCAATTTTGGCTTCAATCAAGGGATACACAGCCACAAT	840	
QY	2781	ATAATACAAGATGTCAGAGGTTTTTACTAGATTTAGACATGCGACATATCAAGATATAC	2840	
DB	841	ATAATACAAGATGTCAGAGGTTTTTACTAGATTTAGACATGCGACATATCAAGATATAC	900	
QY	2841	CAAGCTAAAAAATTAAGTGGTGGTCAAAAAAGAAAGCTGTCAATTAGGAAATTCGTGTTCTT	2900	
DB	901	CAAGCTAAAAAATTAAGTGGTGGTCAAAAAAGAAAGCTGTCAATTAGGAAATTCGTGTTCTT	960	
QY	2901	GGGAAACCAAGATACCTGCTGTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGA	2960	
DB	961	GGGAAACCAAGATACCTGCTGTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGA	1020	

```

RESULT 11
US-10-108-260A-1424
; Sequence 1424, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1

```









Qy	1560	ATTATACAGTTGAAGACCAATGTTTCTCTTTTGAAGAGCTGGAGTCAACTAAAGCTGTTT	1619
Db	1641	ATTATACAGTTGAAGACCAATGTTTCTCTTTTGAAGAGCTGGAGTCAACTAAAGCTGTTT	1700
Qy	1620	ATTATGGAGAAACTGCTGTTGTAGAAATAGATACCTTCCCGAGAGTAATTTTAATA	1679
Db	1701	ATTATGGAGAAACTGCTGTTGTAGAAATAGATACCTTCCCGAGAGTAATTTTAATA	1760
Qy	1680	TACCTAGTTATAGCATTTTCACCTTTTGGATACCTTTTGGCAATTCATATCGTAGCAG	1737
Db	1761	TACCTAGTTATAGCATTTTCACCTTTTGGATACCTTTTGGCAATTCATATCGTAGCAG	1818
RESULT 13			
US-09-822-846-97			
; Sequence 97, Application US/09822846			
; Publication No. US20030027139A1			
; GENERAL INFORMATION:			
; APPLICANT: Jacobs, Kenneth			
; APPLICANT: McCoy, John M.			
; APPLICANT: LaVallie, Edward R.			
; APPLICANT: Collins-Racie, Lisa A.			
; APPLICANT: Evans, Cheryl			
; APPLICANT: Merberg, David			
; APPLICANT: Treacy, Maurice			
; APPLICANT: Agostino, Michael J.			
; APPLICANT: Steining II, Robert J.			
; APPLICANT: Bowman, Michael R.			
; APPLICANT: Spaulding, Vikki			
; APPLICANT: Wong, Gordon G.			
; APPLICANT: Clark, Hilary			
; APPLICANT: Fectel, Kim			
; APPLICANT: Howes, Steven H.			
; APPLICANT: Resnick, Richard J.			
; APPLICANT: Gulkota, Kamalakar			
; APPLICANT: Graham, James R.			
; APPLICANT: Genetics Institute, Inc.			
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS			
; FILE REFERENCE: GIN 6400			
; CURRENT APPLICATION NUMBER: US/09/822,846			
; CURRENT FILING DATE: 2001-03-29			
; PRIOR APPLICATION NUMBER: 60/195,605			
; PRIOR FILING DATE: 2000-04-06			
; NUMBER OF SEQ ID NOS: 629			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 97			
; LENGTH: 1548			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-822-846-97			
Query Match 23.7%; Score 1543.8; DB 11; Length 1548;			
Best Local Similarity 99.8%; Pred. No. 1.3e-309;			
Matches 1545; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
Qy	3386	AGAAATTGACCAAGCAGATTATAGTGATTTTACTCAGCAGCCACTGGAGGAAGAAATGGA	3445
Db	1	AGAAATTGACCAAGCAGATTATAGTGATTTTACTCAGCAGCCACTGGAGGAAGAAATGGA	60
Qy	3446	TTCAAAATCTTTTGATGAATGGAAACAGAGCTTACTTATTTCTTCTGAAACCAAGGCTTC	3505
Db	61	TTCAAAATCTTTTGATGAATGGAAACAGAGCTTACTTATTTCTTCTGAAACCAAGGCTTC	120
Qy	3506	TCTAGTCAGACCAATGAGCCTTTTGGAAACACAGATGTATACATAGCAAGTTTCATTT	3565
Db	121	TCTAGTCAGACCAATGAGCCTTTTGGAAACACAGATGTATACATAGCAAGTTTCATTT	180
Qy	3566	CTTTTACCTTGAAACGCGAAAGTAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTT	3625
Db	181	CTTTTACCTTGAAACGCGAAAGTAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTT	240
Qy	3626	TTTCACAGTTTCAGATTTTTTAATGTTTTTGGTTTCATCATCTTTTTAAAAATGCTGTGGTTC	3685





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 18:55:21 ; Search time 8094 Seconds  
(without alignments)  
19593.119 Million cell updates/sec

Title: US-10-005-338B-1

Perfect score: 6525

Sequence: 1 aaatgtgatattttctct.....ttgatcataataagtgaaat 6525

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estopl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_man:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2934.4	45.0	3950	11 AK047188	AK047188 Mus muscu
2	2807	43.0	3831	11 AK046203	AK046203 Mus muscu
3	1209.6	18.5	1529	11 AK034961	AK034961 Mus muscu
4	1182	18.1	1346	11 U66672	U66672 Homo sapien

# ALIGNMENTS

## RESULT 1

AK047188

LOCUS

DEFINITION

AK047188

Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length

enriched library, clone:B930033A02 product:ATP-BINDING CASSETTE

PROTEIN homolog [Homo sapiens], full insert sequence.

AK047188

AK047188.1 GI:26338625

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

PUBMED 10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalisation and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

PUBMED 11042159

5	837	12.8	850	14	CD250953	CD250953 AGENCOURT
6	761.4	11.7	872	13	BQ722933	BQ722933 AGENCOURT
7	760.6	11.7	809	11	BC029426	BC029426 Homo sapi
8	741.4	11.4	773	13	BQ441353	BQ441353 AGENCOURT
9	731	11.2	747	13	BQ686259	BQ686259 UI-CF-DUI
10	722.4	11.1	771	10	BG564439	BG564439 602584385
11	719.4	11.0	793	14	CB241963	CB241963 UI-CF-FNO
12	681.6	10.4	756	10	BG435656	BG435656 602506942
13	680.4	10.4	744	9	AL700140	AL700140 DKFPZp686H
14	676.4	10.4	697	13	BQ774317	BQ774317 UI-H-EZ1-
15	668	10.2	926	13	BUI46839	BUI46839 AGENCOURT
16	661.6	10.1	728	13	BUE23566	BUE23566 UI-H-FG1-
17	655.4	10.0	792	10	BG564921	BG564921 602589884
18	652.4	10.0	800	13	BUI750944	BUI750944 CH3#036 H
19	646	9.9	735	12	BG926885	BG926885 HNC24-1-E
20	641	9.8	664	14	CB044826	CB044826 NISC GC06
21	634.4	9.7	670	10	BG570726	BG570726 602591840
22	633	9.7	633	14	CA773136	CA773136 im60B03.Y
23	631.8	9.7	3890	11	BC053340	BC053340 Mus muscu
24	630.4	9.7	768	13	BUI853701	BUI853701 AGENCOURT
25	628.6	9.6	724	14	CB962917	CB962917 AGENCOURT
26	625.2	9.6	694	10	BF969534	BF969534 602271862
27	623	9.5	623	14	CA773849	CA773849 im58A04.Y
28	617	9.5	642	14	CA773183	CA773183 im60G04.Y
29	612.4	9.4	614	14	CA771132	CA771132 io71b12.X
30	610.8	9.4	3357	11	BC037309	BC037309 Homo sapi
31	610.2	9.4	4119	11	AK029256	AK029256 Mus muscu
32	589.6	9.0	898	10	BGI92651	BGI92651 RST11768
33	586.8	9.0	961	14	CD242338	CD242338 AGENCOURT
34	583.8	8.9	589	10	BGI49983	BGI49983 nad50a05.
35	583.6	8.9	678	10	BE220217	BE220217 hv69f11.X
36	573.6	8.8	581	14	CA941420	CA941420 ir34h12.X
37	573.2	8.8	578	13	BU069128	BU069128 im58A04.X
38	572.4	8.8	629	14	CA771439	CA771439 io71b12.Y
39	565.8	8.7	608	9	AW131017	AW131017 xe68f06.X
40	561.6	8.6	570	2	HSM090776	Bx497427 Homo sapi
41	561.2	8.6	622	13	BU952395	BU952395 io74a05.X
42	559.4	8.6	769	14	CD349566	CD349566 UI-M-FY0-
43	557.6	8.5	3300	11	AK085555	AK085555 Mus muscu
44	553.2	8.5	571	14	W63628	W63628 zc56f10.r1
45	547.6	8.4	639	12	BG928068	BG928068 HNC16-1-A

REFERENCE AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861	source	1. 3950 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:B930033A02" /db_xref="taxon:10090" /clone="B930033A02" /tissue_type="cerebellum" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="10 days neonate" 201. ->3950 /notes="unnamed protein product; ATP-BINDING CASSETTE PROTEIN homolog [Homo sapiens] (SPTP[CAB93535], evidence: FASTA, 89.4%ID, 76.5%length, match=3750) putative" /codon_start=1 /protein_id="BAC32984.1" /db_xref="GI:26338626" /translation="MATAIRDVGVWRQTRTLKKNYLIKRTKSSVQOEILPFLPFLF WLIVSMHNNKVEEVDIELSPDKFSLSNVLGTYTPVTNITSSIMQRCVTDHLPK VIVTEYANKELVAASLSKSNFVGVEEDTMSYELRFPPEMIPVSSIYVNSREGCS KTCDAQWSLGFVLAQSIDAAIIQLKTNVSVWSELESTKAVIMGAAVVEIDTFR GVILYIVAFSPFGYFLAIHVAEKKLKEFKIMGLDHTAFWLSWLLIYASLFL MSLMVATATASSLFPSSSIVIFLFLFYGLSSVFFALMLTLPFKSKHGVVVEFF TVFVGGLIIVLIESPRLWMLFSPQLCQAFLLGIAQVHLEDFNEGLPFLVLTGEG PYPLIITIMLALDSVFYLLAVYLDQVIGFGRSLYFLKPSYVNSKNRYKEL SEGNNINISLNEIVEPVSEFFIGKEAIRISGIOKSYRKTENVEALRNLSDIYEGQ ITALHSGTGKSTLMNLGCLCPSPGPAISYGHVSEIDEMEAEKMGICQPSDI NFDVLTVEENSLILASIKGPANNIIQEVQKVLDDMQAIIKQNAKKLSGGQKKLS GVAVLGNPKILLIDBPTAGMDPCSRHIVNLLKRYKANRVTVSTHFMDEADILLADR KAVISQMLKCVGSSIFLKSQWIGYRLSMYIDRYCATESSLSLVROHIFAAALQON DQVLVSLPFKDMDFSLGFSALDHSNLGVISYGVMTLEDVFLKLEVAEIDQAD YSVTQPREEDTSKFSDEMEQSLILSETKASSVSTMSLWKQCVSTIAKFHLSLK RESKVRALLLLIIFAVQIFMFLHSPKNAVVPKLPVLDLYFLKPGDKPKYKTS LLQNSVDSINGLIEFEAHQNTVMAMPNDSDYVSAAPHASALNVRSEKDVVESA NSTWYCLPVMMNIIISNYLYHLNVTALQIOTWSTPFIQIOTIDVFKLELYFOAALLGI IVTAMPPYFAMENAEHNIKAYTQLKSLGPSAYVWGQVVDIPLFFVLILMLGSL FAPHGLVFPKFLAVVFCILAYVPSVILTYIASFTFKKLTKEFWMSFYVSTAL ACVATETTFLOYAVTAVHYTFCIAPIYPLLAGCLISFIKSGSKMKNENTYNPW DRLVAVIMPYLQCLIMFILLQHYEKIHGGRSIRKDPFRALSQKA"
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wyshak-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001) 21085660 11217851	CDS	BASE COUNT 1055 a 802 c 863 g 1230 t ORIGIN Query Match 45.0%; Score 2934.4; DB 11; Length 3950; Best Local Similarity 86.2%; Pred. No. 0; Matches 3247; Conservative 0; Mismatches 522; Indels 0; Gaps 0; QY 991 AATAGGTTTATTCAGAAACATGCTCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGA 1050 DB 181 AGTAACTGACCTAGAAACATGCTACTGCTCAATTAGGAGTGTGGAGTGTGGACAGA 240 QY 1051 CCAGAACATCTTACTGAGAAATTTACTTAATTAATGAGAACCAACCAAGAGTAGTGTC 1110 DB 241 CCAGAACATCTTACTGAGAAATTTACTTAATTAATGAGAACCAACCAAGAGTAGTGTC 300 QY 1111 AGGAAATCTTTTCCACTATTTTATTTTGGTTTATTAATTAATGAGTATGTCATC 1170 DB 301 AGGAAATCTTTTCCCTCTATTTTCTTATTTTGGCTGATTAATGAGTATGTCATC 360 QY 1171 CAATAAGAAATATGAGAGAGTGCCTAATATAGAACTCAATCTTATGAGCAAGTTTACTC 1230 DB 361 CAATAAGAAATATGAGAGAGTGTCTGATATAGAGTTCAGCCCTATGACAAATTCAGCC 420 QY 1231 TTTCTAATCTAATCTTCGGATATCTCCAGTGACTAATATTTACAGAGAGCATGCGAGA 1290 DB 421 TTTCCAGGTTATTTCTTGGATACACTCCCGTGACTAACATTTACAGAGAGCATGCGAGA 480 QY 1291 AAGTGTCTACTGATCATCTACTGATGTCTAATAATTTACTGAAGAAATATACAAATGAAAAG 1350 DB 481 GGGTTTGACGATCACTTCCCAAGGTATAGTTACTGAAGAAATACGCAATGAGAAAG 540 QY 1351 AAATGTTAACATCCAGTCTCTCTAAGCCGAGCAACTTTGTAGGTGTGTTTCAAGACT 1410
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 3950) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers	FEATURES	



3571 CCTTGAACGTGAAGTAATCAGTCAGATCAGTGTGCTTCTGCTTTTAAATTTTTTCA 3630  
Db  
2761 CATTTGAAACGAGAAAGCAATCAGTGGCGCTGTGCTTCTGCTTTTAAATTTTTTG 2820  
Qy  
3631 CAGTTGAGATTTTATGTTTGGTTTCATCAGCTCTTTTAAAAATGCTGTGGTTCCCATCA 3690  
Db  
2821 CAGTTGAGATTTTATGTTTCTCCATCATCTTTTAAAAATGCTGTGGTTCCCATCA 2880  
Qy  
3691 AACTTGTTCAGACTTATATTTCTAAACCTGGAGACAACACACATAATACAAAACA 3750  
Db  
2881 AACTCGTTCAGACTTGTATTTCTTAAAGCTGGAGATAAACCTATAATACAAAACA 2940  
Qy  
3751 GTCTGCTCTTCAAAATTCGTGACTCAGATCAGTGAATCTTATAGCTTTTTCACAA 3810  
Db  
2941 GCTGCTGCTTCAAAATTCGTGACTCAGATCAATGCTTATGAGTTTGTGAC 3000  
Qy  
3811 GCAGAACATAATGGTGACGATGATTAATGACAGTCACTATGTCGTTGCGTTCCTCAT 3870  
Db  
3001 ACCAGAACATAATGGTGCAATGTTTAAATGACAGTCACTATGTCGTCGCTCCTCA 3060  
Qy  
3871 GTGCGGCTTTAAATGATGATCATTAGAAAGGACTATGTTTGGAGCTGTTTTCACA 3930  
Db  
3061 GTGCGGCTCTCAATGTGGTGGTCTTCTGAAAGGACTATGTTTCTGCTGTTTTCACA 3120  
Qy  
3931 GTACTATGTTTATCTTTACCTATATAGTGAATCACTAGTAACTACTATCTTTATC 3990  
Db  
3121 GTACTATGTTTATGTTTCCAGTCATGATGAACATCATAGTAACTACTATCTTTATC 3180  
Qy  
3991 ATTTAAATGTGACTGAACACCATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATTA 4050  
Db  
3181 ATTTAAATGTGACTGAAGCCATCCAGACTGGAGTACCCCGTTCAITTCAGAAATTA 3240  
Qy  
4051 ATATAGTTTAAATGTGACTGTATTTTCAAGCAGCTTGTGTAATCATTTGTTACTG 4110  
Db  
3241 ACATTTGTTTAAATGTGACTGTATTTTCAAGCAGCTTGTGTAATCATTTGTTACTG 3300  
Qy  
4111 CAATGCCACCTTACTTTGGCCATGGAATGACAGAAATGACAGAAATCAAGCTTATATC 4170  
Db  
3301 CAATGCCACCTTACTTTGGCCATGGAATGACAGAAATCAAGCTTATATC 3360  
Qy  
4171 AACTTAACTTTCAGGTCCTTTTGGCCATGTCATATTTGGAATGACAGCTGTTGTTGATA 4230  
Db  
3361 AACTTAACTTTCAGGTCCTTTTGGCCATGTCATATTTGGAATGACAGCTGTTGTTGATA 3420  
Qy  
4231 TCCCTTATTTTATCATTTCTTATTTTGTAGTGTAGGAGCTTACTGGCATTCATATG 4290  
Db  
3421 TCCCTTATTTTATCATTTCTGATTTTGTAGTGTAGGAGCTTATTTGCAATTCATCATG 3480  
Qy  
4291 GATTATATTTTATGTAAGTTCCTGCTGTTGTTTTCCTTATTTGTTGTTATGTTTC 4350  
Db  
3481 GACTGTATTTTATCTGCAAAATTTCTGCTGTGTTGTTTTCCTTATTTGCTTATGTTGC 3540  
Qy  
4351 CATCAGTATTCCTGTCATTTATGCTTTCTTTTCACTTTTAAAGAAATTTTAAATACCA 4410  
Db  
3541 CCTCCGTCATTCCTGTCAGTACATAGCTTCGTTCACTTTCAAGAAATTTTAAATACCA 3600  
Qy  
4411 AAGAAATTTGGTCATTTATCTATCTGTCGAGGTTGCTTGTATGTAATCACTGAA 4470  
Db  
3601 AGAAATTTGGTCATTTATCTATCTGTCGAGGCTTGTGTCGCAATCACCGAA 3660  
Qy  
4471 TAACTTTCTTTATGGGATACAAATGCAACTATCTTCAATGCTTTTGTATCATCA 4530  
Db  
3661 CACTTTCTTTCTGCAATATGCAATGCTGCTGTTTCTTCAATACCTTCTGATGACCA 3720  
Qy  
4531 TTCAATCTATCCACTTCTAGGTTGCTGATTTCTTTTCAATAAGATTTCTTGAAGAATG 4590  
Db  
3721 TTCAATCTATCCCTCTCTCTGTTGCTGATTTCTTTTCAATAAGGTTCTTGAAGAATA 3780  
Qy  
4591 TAGCAAAATATGGACACCTATATCCATGGATAGGCTTTTCACTAGCTGTTATATCGC 4650  
Db  
3781 TGCAAAATATGAAATACCTTACAACTCCCTGGATAGACTTTTGTGCTGTATCATG 3840  
Qy  
4651 CTTACCTGCAGTGTGCTGCTGATTTTCTCTTACAAATCTATGAGAAAAATATGGAG 4710

3841 CGTACCTCAGTGTATCTGTTGATTTTCTCTTACAACTATGAGAAATACATGAG 3900  
Qy  
4711 GCAGATCAATGAAGAAAGATCCCTTTTTCAGAAACCTTTCAACGAAGTC 4759  
Db  
3901 GCAGATCTATAAGAAAGATCCATTTTTCAGGGCCCTTTTCAAAAAGGC 3949  
RESULT 2  
AK046203 3831 bp mRNA linear HTC 05-DEC-2002  
LOCUS Mus musculus adult male corpora quadrigemina cDNA, RIKEN  
DEFINITION full-length enriched library, clone:B230352G14 product:ATP-BINDING  
CASETTE PROTEIN homolog [Homo sapiens], full insert sequence.  
ACCESSION AK046203  
VERSION AK046203.1 GI:26091381  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1  
AUTHORS Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE 2  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
REFERENCE 3  
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kasai, H.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
REFERENCE 4  
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, I.,  
Sato, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Kado, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Mateu, Y., Nikaide, I., Pesole, G.,  
Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,  
Wagner, L., Watanabe, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohtsuki, S.  
and Hayashizaki, Y.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409 (6821), 685-690 (2001)  
MEDLINE 21085660  
PUBMED 11217851  
REFERENCE 5



**AUTHORS** The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

**TITLE** Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

**JOURNAL** Nature 420, 563-573 (2002)

**REFERENCE** 6 (bases 1 to 3831)

**AUTHORS** Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tonaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

**TITLE** Direct Submission

**JOURNAL** Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

**COMMENT** cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/

**FEATURES** Location/Qualifiers

1..3831

organism="Mus musculus"

mol\_type="mRNA"

strain="C57BL/6J"

db\_xref="FANTOM DB:B230352G14"

db\_xref="taxon:10090"

clones="B230352G14"

sex="male"

tissue type="corpora quadrigemina"

clone lib="RIKEN full-length enriched mouse cDNA library"

dev stage="adult"

misc\_feature 1..3831

note="ATP-BINDING CASSETTE PROTEIN homolog [Homo sapiens] (SPTR|CAB93535, evidence: FASTA, 89.4%ID, 76.5%length, match=3750)"

BASE COUNT 1021 a 767 c 838 g 1205 t

ORIGIN

Query Match 43.0%; Score 2807; DB 11; Length 3831;

Best Local Similarity 86.3%; Pred. No. 0;

Matches 3112; Conservative 0; Mismatches 491; Indels 1; Gaps 1;

QY 991 AATAGGTTTATTGAGAAACATGTCCTCACTGCAATTAGGAGGTAGGAGTTGGAGACAGA 1050

DB 227 ATTAAACTGACCTAGAGAAACATGGCTACTGCAATTAGGAGGTAGGAGTTGGAGACAGA 286

QY 1051 CCAGAACACTTCTACTGAAAGATTACTTAATTAATGCAGAACCAAAAGAGTAGTGTC 1110

DB 287 CCAGAACACTTCTACTGAAAGATTACTTAATTAATGCAGAACCAAAAGTAGTGTC 346

QY 1111 AGGAAATCTTTTCCACTATTTTATTTTGGTTTAAATTAATTAATGATGATGATC 1170

DB 347 AGGAAATCTTTTCCCTATTTTCTTATTTTGGCTGATATTAGTTAGCATGATGATC 406

QY 1171 CAATAGAAATATGAGAGAGTGCCTAATATAGNACTCAATCCTATGGAAGTTTACT 1230

DB 407 CAAATAGAAATATGAGAGAGTATCTGATATAGAGCTCAGCCCTATGGAAGTTTACTGAC 466

QY 1231 TTCTTAATCTAATCTTCTTGGATATATCTCCAGTGACTAATATTACAAGCAGCATCATGCGA 1290

DB 467 TTTCACACGTTATTTCTTGGATACACTCCCGTGACTAACATTACAAGCAGCATTATGCGA 526

QY 1291 AAGTGCTACTGATCATCTACCTGATGTCATATTAATCTAGGAATATACAAATGAAAG 1350

DB 527 GGGTTTCTACCGATCATCTTCCCAAGGTATAGTTACTGAAGATATAGCAATAGAGAAAG 586

QY 1351 AAATGTTTAAACATCCAGTCTCTCTAAGCCGACCAACTTTGTAGGTGTGGTTTCCAAAGACT 1410

DB 587 AACTGGTAGCCGCAAGCTCTTCTAAGTCAGCAACTTCGTAGGTGTGGTTTCAAAGACA 646

QY 1411 CCAATGCTTATGAACCTGGTTTTCCTGATATGATTCAGATATCTTCTATTTATATGG 1470

DB 647 CCAATGCTTATGAACCTGGTTTTCCTGATATGATTCAGATATCTTCTATTTATATGA 706

QY 1471 ATTCAAGAGCTGCTCTTCAAAATCATGTGAGGCTGCTCAGTACTGCTCTCAGTTTCA 1530

DB 707 ATTCAAGAGAGAGGCTGTTCAGAGACATGTGATGCTGCTCAGTACTGCTCTTGGGGTTA 766

QY 1531 CAGTGTTCACAGCATCCATAGATGCTGCCAATTATACAGTTTGAAGACCAATGTTTCTCTTT 1590

DB 767 CAGTGTTCAGGCATCGATAGATGCTGCCAATTATACAGCTGAAGACCAATGTTTCTGT 826

QY 1591 GGAAGAGCTGGAGTCAACTTAAAGCTGTATTATATGGAGAAACTGCTGTTGTAGAAATAG 1650

DB 827 GGAGCAGCTGGAGTGCACCAAGCTGTGATCATGGAGAGGCGCTGTTGTGGAGATTG 886

QY 1651 ATACCTTTCCCGAGGAGTAATTTTAAATATACCTAGTTATAGCATTTTACCTTTTGAT 1710

DB 887 ACACCTTCCCGAGGAGGCTCATCTCATCTACCTCGTCATAGGCTTCTCCGCCCTCGGCT 946

QY 1711 ACTTTTGGCAATTCATATCGTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGA 1770

DB 947 ACTTCTGGCAATCCATCGTGGCAGAAAAAGAAAGAGTTAAGAAATTTTAAAGA 1006

QY 1771 TAAATGGAGCTTCAATGATCTGCTTTTGGCTTTCTGGGTTCTCTATATACAAAGTTAA 1830

DB 1007 TAAATGGAGCTTCAATGATCTGCTTTTGGCTTTCTGGGTTCTCTGACCAAGCTTGA 1066

QY 1831 TTTTCTTATGCTCCCTCTTATGGCAGTCAATGGCAGAGCTTCTTGTGTTATTTCTCTCAA 1890

DB 1067 TTTTCTTATGCTCCCTCTTATGGCTGTCTATGCTCAACAGCTTCTCTGTTATTCCTCAGA 1126

QY 1891 GTAGCAGCATTTGATATTTCTGCTTTTCTTTTATGATGATATCATCTGTATTTTGTG 1950

DB 1127 GTAGCAGCATTTGATATTTCTTCTACTGTTATCTTATGAGGATGTCACTGTGTTTGT 1186

QY 1951 CTTTAAATGTCGACACCTCTTTTAAATAATCAAAACATGTGGGAATAGTTGAAATTTTGTG 2010

DB 1187 CTTTAAATGTCGACGCTCTTTTAAATAATCAAAACATGTGGGAGTGTGTTGAGTTTGTG 1246

QY 2011 TTAATGCTGGCTTTTGGATTTATTTGGCTTTATGATTAATCTCATAGAAAGTTTTCCTCAAT 2070

DB 1247 TCACCGTGGTGTGTTGATTTGTTGGCTGTCTGATTTGCTCTCATAGAAAGTTTCCCCAGGT 1306

QY 2071 CGTTAGTGTGGCTTTTTCAGTCTTTCTGTCACGTACTTTTGTGATTTGTTGATTTGTCACAG 2130

DB 1307 CGTGTGTGGCTTCTTCAGTCTTTTGTGTCAGTGTGCTTTCTGATTTGGGATTTGCACAG 1366

QY 2131 TCATGATTTTGAAGATTTTAAATGAAGGTGCTTCAATTTTCAAAATTTGACATGACGGCCAT 2190

DB 1367 TCATGATTTTGAAGATTTTAAAGAGGTGCTTATTTTCTAATTTGACTGAAGGTCCCT 1426

QY 2191 ATCCTCTAATTAATTAACAATATCATGCTCACTTAATAGTATATTTCTATGCTCTCTGG 2250

DB 1427 ATCCTCTAATTAATTAATATCATGCTAGCTCTTGCACAGTGTGTTCTATGCTCTCTGG 1486

QY 2251 CTGCTCTCTTGTATCAAGTCAATTCAGGGGAAATTTGCTTTACGGAGATCATCTTATATT 2310

DB 1487 CTGCTCTCTGACCAAGTCAATTCAGGGGAAATTTGCTTTGAGGAGGTCACTTTGTATT 1546

QY 2311 TTCTGAAGCTTCTCATATTGCTC-AAAGAGTAAAAAGAAATTTATGAGGAGTTTATCAGAGGC 2369

1547 TTTTGAAGCCATCGTATGGTCAAAACAAACAAAGAGCTATCGAGGCGC 1606  
2370 AATGTTAAATGGAATATAGTTTATAGTGAATATATAGCCAGTTCCTTCAGAAATTTGTA 2429  
1607 AACATTAATGGCAATATAGTCTCAATGAATTTGTTAGCCCGCTTCTTCAGAAATTTATA 1666  
2430 GGAAGAGAGCCATAGAAATAGTGTATTCAGAACACATACAGAAAGAGGTGAATAT 2489  
1667 GGAAGAGAGCTATAGAAATAGTGTATTCAGAAATCTCTATAGAAAGAAACTGAGAAC 1726  
2490 GTGGAGGCTTTGAGAAATTTGTCATTTGACATATATAGGGTCAGATTACTGCTTACTTT 2549  
1727 GTGGAGGCTTTGAGAAATTTGTCATTTGACATATATAGAGTCAGATTACTGCACTGCTG 1786  
2550 GGCCACAGTGAACACAGAAAGAGTACATTTGATGAATATTTCTTTGTGGACTCTGCCACCT 2609  
1787 GGCCACAGTGAACACAGAAAGAGTACATTTGATGAATATTTCTGTGGACTGTGTCCACCC 1846  
2610 TCTGATGGGTTGCATCTATATATGACACAGAGTCTCGAAATAGATGAATATTTGAA 2669  
1847 TCTGATGGGTTGCTCTATATATGACACAGAGTCTCTGAAATAGATGAATATTTGAA 1906  
2670 GCAAGAAATATGATGGCATTTGTCACACAGTTAGATATACATTTGATGTTTGCAGTA 2729  
1907 GCAAGAAATATGATGGCATATGTCGCGAGTCAGATATAAATTTGATGTTCTGCAGTA 1966  
2730 GAAGAAATTTATCAATTTTGGCTTCAATCAAGAGGATACCGCCCAACATATATACAA 2789  
1967 GAAGAAATTTATCAATTTTGGCTTCAATCAAGAGGATACCGCCCAACATATATCA 2026  
2790 GAAGTGCAGAGGTTTACTAGATTTAGACATGACAGTATCAAGATACCAAGCTTAA 2849  
2027 GAAGTGCAGAGGTTTACTAGATTTAGACATGACAGTATCAAGATACCAAGCTTAA 2086  
2850 AATTAAGTGGTGGTCAAAAGAAAGCTGTCTAATAGGAATGCTCTTCTTGGGAACCCA 2909  
2087 AAGTTAAGCGGTGGTCAGAAAGAGGAGTGTCTGTAGGAATTTGCACTTCTCGGAATCCA 2146  
2910 AAGATAGTGTCTAGATGAACCAACAGTGGAAATGGACCCCTGTTCTCGACATATGTA 2969  
2147 AAGATAGTGTCTAGACGAGCTTACAGCAGGAATGGACCCCTGCTCTCGCCATATTTGT 2206  
2970 TGGAACTTTTAAATACAGAAAGCAATCGGTCAGAGTGTCTAGTACTCATTTTCATG 3029  
2207 TGGAACTTTTAAAGTATAGAAAGGCTTACAGAGTGCACCGTGTGTAGTACTCTCTCAIG 2266  
3030 GATGAAGCTGACATTTCTTGAGATAGGAAGCTGTGATATCAAGGAATGCTGAAATGT 3089  
2267 GATGAAGCTGACATTTCTTGCGACAGGAAGCTGTCTATATCAAGGAATGCTGAAATGT 2326  
3090 GTTGGTTCCTCAATGTTCTCAAAAGTAATAGGGGATCGGCTACCGCTGAGCATGTAC 3149  
2327 GTTGGTTCCTCAATTTTCTTAAAGAGTAAATGGGAATCGGCTACCGCTGAGCATGTAT 2386  
3150 ATAGACAAATTTGTCGACAGATCTCTTCTTCACTGCTGTTAAACAAATATACCTGGA 3209  
2387 ATAGACAGATAGTGTGCGACAGAGTCACTGTGCTGCTGTGTAGGAGACATATCCCGCA 2446  
3210 GCTACTTTTATACAAAGAGATGACCAACATTTGTGTATAGCTTGTGCTTTCAAGGACATG 3269  
2447 GCCCGCTACTGACAGAGATGACCAAGCTGTGTGACAGCTGCTTTCAGAGACATG 2506  
3270 GACAAATTTTCAAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3329  
2507 GACAAATTTTCAAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2566  
3330 TATGCTGCTTTCATGACGCTTTGGAGAGCTATTTTAAAGCTAGAGTGTGAGCAGAA 3389  
2567 TATGCTGCTTTCATGACGCTTTGGAGAGCTATTTTAAAGCTAGAGTGTGAGCAGAA 2626  
3390 ATTGACCAAGCAGATATAGTGTATTTTACCTGACAGCCACTGGAGGAAGAAATGGATTCA 3449

2627 ATTGACCAAGCAGACTATAGTGTATTTTACAGAGCGCGGAGGAGAAACAGATTCA 2686  
3450 AAATCTTTTGTATGAAATGGAACAGAGCTTACTTATCTTTCTGAAACCAAGGCTTCTCTA 3509  
2687 AAATCTTTTGTATGAAATGGAACAGAGTCTTACTTATCTCTGAAACCAAGGCTTCTCTA 2746  
3510 GTGAGCACCATTGAGCCTTTTGGAAACAAACAGATGATACAAATGAGCAAGTTTCTTTT 3569  
2747 GTGAGCACCATTGAGCCTTTGGAAGCAGCAAGTGTCTAGATTGCAAAAGTTTCAATTTCTC 2806  
3570 ACCTTGAACCTGGAAGTAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTCTTTC 3629  
2807 TCAATGAAACGAGAAAGCAATCAGTGGCGCTGTGTGCTTCTGCTTTTAAATTTTCTT 2866  
3630 ACAGTTCAGATTTTATGCTTTTGTGCTCATCACTCTTTTAAATATGCTGTGCTTCCCATC 3689  
2867 GCAGTTCAGATTTTATGCTTTTGTGCTCATCACTCTTTTAAATATGCTGTGCTTCCCATC 2926  
3690 AAATCTTTTCCAGACTTATATTTTCTAAACCTGGAGCAACCAACATATAAATAACAAACA 3749  
2927 AAATCTCGTTCAGACTTGTATTTCTTAAAGCCTGGAGATAAACCTCATATAAATAACAAACA 2986  
3750 AGTCTGCTTCTTCAAAATCTGCTGACTCAGATATCAGTGTATCTTATTTAGCTTTTCTACA 3809  
2987 AGCTGCTGCTTCAAAATCTTACTGACTCAGATATCAATGGTCTTATTTAGTGTTTTGTGCA 3046  
3810 ACCAGAAACATAATGGTGAACGATTAATGACAGTACTATGATCGTGGCTCCCCAT 3869  
3047 CACAGAAACATAATGGTGGCAATGTTTAAAGCAGTACTATGCTGTGCTGCTTCTCAC 3106  
3870 AGTGGGCTTTTAAATGCTGATGCATTTAGAAAGGACTATGTTTTTGGAGCTGTTTTCAAC 3929  
3107 AGTGGGCTCTCAATGCTGCTGCTTCTGAAAGGACTATGTTTTTCTGCTGTTTTCAAC 3166  
3930 AGTACTATGTTTATCTTTTACCTATATTTAGTGAATATCATTAGTACTATCTTATTTAT 3989  
3167 AGTACTATGTTTATGTTTGGCAGTCAATGATGAACATCATTTAGTACTATCTTATTTAT 3226  
3990 CATTTAAATGCTGAAACCATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTACT 4049  
3227 CATTTAAATGCTGAAAGCCATCCAGACTTGGAGTACCCCGTTTCAATTTCAAGAAATTACT 3286  
4050 GATATAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACT 4109  
3287 GACATTTGTTTTTAAATTTGAGCTATATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACT 3346  
4110 GCAATGCCACCTTACTTTGCCATGGAATGCAAGATCATAGATCAAGGCTTTTACT 4169  
3347 GCAATGCCACCTTACTTTGCCATGGAATGCAAGATCATAGATCAAGGCTTTTACT 3406  
4170 CAACTTAACTTTTCAAGGCTTTTGGCATCTGCATATTCGATTTGGACAGCTTTGTTGAT 4229  
3407 CAATTTAACTTTTGGGCTTTTGGCATCAGCTTATGGTGGACAGCTTTGTTGAT 3466  
4230 ATCCCTTTTATTTTATCATTTCTTATTTTGTGCTAGGAAGCTTACTGCGCATTTCTAT 4289  
3467 ATTCCTTTTATTTTGTGTTCTGATTTTGTGCTGGAAGTTTATTTGCACTTTTCTCAT 3526  
4290 GGAATATATTTTATATATGTAAGTTCCTGCTGCTGTTTTTGTGCTTATTTGCTTATGTT 4349  
3527 GGAATATATTTTATATATGTAAGTTCCTGCTGCTGTTTTTGTGCTTATTTGCTTATGTT 3586  
4350 CCATCAGTATTTCTGCTTCACTTATTTGCTTCTTCTTCACTTTTAAAGAAATTTTAAATACC 4409  
3587 CCCTCGCTTCTGCTTCACTTATTTGCTTCTTCTTCACTTTTAAAGAAATTTTAAATACC 3646  
4410 AAAGAAATTTGCTTCACTTATTTGCTTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 4469  
3647 AAGGAATTTGCTTCACTTATTTGCTTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 3706  
4470 ATAACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4529  
3707 ACAACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3766



```
QY 3405 TATAGTGTATTTACTCAGCAGCCACTGGAGGAAAGAAATGGAATCAAAATCTTTTGATGAA 3464
Db |||||
QY 1 TATAGTGTATTTACACAGCAGCGCGGAGGAGAAACAGATTCAAAATCTTTTGATGAA 60
Db |||||
QY 3465 ATGGAACAGAGCTTACTTATTTCTTTCTGAAACCAAGGCTTCTTAGTGAGCACCATGAGC 3524
Db |||||
QY 61 ATGGAACAGAGCTTACTTATTTCTTTCTGAAACCAAGGCTTCTTAGTGAGCACCATGAGC 120
Db |||||
QY 3525 CTTTGGAAACAAACAGATGTATACAAATAGCAAGCTTTTCATTTCTTACCTTGGAAACCTGAA 3584
Db |||||
QY 121 CTCTGGAAGCAGCAAGGTCTACGATTCGCAAGTTTCATTTCTCTCATTTGAAACGAGAA 180
Db |||||
QY 3585 AGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTTT 3644
Db |||||
QY 181 AGCAATCAGTGGCGCTGTGCTTCTGCTTTTAAATTTTTCGAGTTTCAGATTTTTT 240
Db |||||
QY 3645 ATGTTTTGTTTCATCATCTTTTAAATGCTGTGTTCCCATCAAACTTTGTTCCAGAC 3704
Db |||||
QY 241 ATGTTTTTCTCCATCATTTCTTTTAAATGCTGTGTTCCCATCAAACTTCGTTCCAGAC 380
Db |||||
QY 3705 TTATATTTCTTAAACCTGGAGACCAACCATATATACAAACAAAGTCTGCTTCTTCAA 3764
Db |||||
QY 301 TTGTATTTCTTAAAGCCTGGAGATAAACCTCATATAATACAAACAAAGCCTGCTGCTCAA 360
Db |||||
QY 3765 AATTCTGCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAGCCAGAACATAATG 3824
Db |||||
QY 361 AATTCTACTGCTCAGATATCAATGCTTTATTAGTGTATTTTTCACACACAGAACATAATG 420
Db |||||
QY 3825 GTGACGATGATTAATGACAGTACTATGATCCGCTGCTCCCATAGTGGCGCTTTAAAT 3884
Db |||||
QY 421 GTGCAATGTTTAAATGACAGTACTATGCTGCTGCTCCTCACAGTGGCGCTCTCAAT 480
Db |||||
QY 3885 GTGATGCAATTCAGAAAGGACTATGTTTTTTCAGCTGTTTTTCAACAGTACTATGTTTTAT 3944
Db |||||
QY 481 GTGGTGGCTTTCTGAAAGGACTATGTTTTTCTGCTGTTTTTCAACAGTACTATGTTTTAT 540
Db |||||
QY 3945 TCTTTTACTATATTAGTGAATATCATTAGTAACCTACTATCTTTATCAATTTAAATGTGACT 4004
Db |||||
QY 541 TGTTTGGCAGTCATGATGAACATCATTTAGTAACCTACTATCTTTATCATTTAAATGTGACT 600
Db |||||
QY 4005 GAAACCATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAATCTGATATAGTTTTTAA 4064
Db |||||
QY 601 GAAGCCATCCAGACCTGGAGTACCCCGTTTCAATCAAGAAATTAATCTGACATTTGTTTTAA 660
Db |||||
QY 4065 ATTGAGCTGATTTTCAAGCAGCTTTGCTGGATCATTTGTTACTGCAATGCCACCTTAC 4124
Db |||||
QY 661 ATTGAGCTATATTTTCAAGCAGCTTTGCTGGATCATTTGTTACTGCAATGCCACCTTAC 720
Db |||||
QY 4125 TTTGCCATGAAATGACAGAGATCAATAGATCAAGCTTATATCTCAACTTAACTTTTCA 4184
Db |||||
QY 721 TTTGCCATGAAATGACAGAGATCATAGATCAAGCTTATATCTCAACTTAACTTTTGG 780
Db |||||
QY 4185 GGTCTTTTGGCATCTGCATATGGATGGACAGCTGTTGTTGATATCCCTTATTTTTT 4244
Db |||||
QY 781 GGTCTTTTGGCATCAGCCTATGGTGGACAGCTGTTGTTGATATCCCTTATTTTTT 840
Db |||||
QY 4245 ATCATTTCTATTTTGTAGCTAGGAGCTTACTGGCATTTTCATTTGATATATTTTTAT 4304
Db |||||
QY 841 GTTGTCTGATTTTGTAGCTGGGAAGTTTATTGTCATTTTCATGAGCTGTTATTTTTAT 900
Db |||||
QY 4305 ACTGTAAGTTTCTTGTGCTGGTTTTTTTGGCTTATTGTTGTTATGTTCCATCAGTTATCTG 4364
Db |||||
QY 901 CTTGCAAAATTTCTTGTGCTGGTTTTTGGCTCATTTGCTTATGTCCTCCGTCATTTCTG 960
Db |||||
QY 4365 TTCATTTATATGCTTTCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCA 4424
Db |||||
QY 961 TTCAGGTACATAGCTTCTGTTTCACTTTTCAAGAAATTTTAAATACCAAGAAATTTTGGTCA 1020
Db |||||
QY 4425 TTTATCTATTTCTGCGCAGGTTGCTTTGTTATGCAATCACTGAAATTAACCTTTCTTTATG 4484
Db |||||
QY 1021 TTTATCTATTTCTGTCGCGCATTTGGCTTTGTGTCGAATCAACGAAACAACTTTCTTTCTG 1080
Db |||||
QY 4485 GGATACAAATTTGCAACTATTTCTTCTTATTATGCTTTTGTATCATCATTTCCAAATCTATCCA 4544
Db |||||
```

```
Db |||||
QY 1081 CAATATGCAGTTACGGCTGTCTTTTACATTTACACCTTCTGCTAGCCATTTCCAATCTACCC 1140
Db |||||
QY 4545 CTTCTAGGTTCCCTGATTTCTTTTTCATTAAGATTTCTTTGGAAGAATGTACGAAAAAATCTG 4604
Db |||||
QY 1141 CTCCTGGGTTGTCGATTTCTTTTTCATTAAGGGTTCTTTGGAAGAATATGCCCCAAAAATGAG 1200
Db |||||
QY 4605 GACACCTATAATCCATGGATAGGCTTTTTCAGTAGCTGTTTATATGCTTACCTGCTGAGTGT 4664
Db |||||
QY 1201 AATACCTACATCTCTGGATAGACTTTTGTGCTGTAATCATGCTCCCTACCTGAGTGT 1260
Db |||||
QY 4665 GTACTGTGGATTTTCTCTTCAATACTATGAGAAAAAATATGGAGGAGATCAATAAGA 4724
Db |||||
QY 1261 ATACTGTGGATTTTCTCTTACAACTATGAGAAAAATACATGGAGGAGATCTATAAGA 1320
Db |||||
QY 4725 AAGATCCCTTTTTCAGAAAACTTTTCAACGAAGTCTAAAAATAGGAAGCTTCCGAACCA 4784
Db |||||
QY 1321 AAGGATCCATTTTTCAGGGCTTTTTCACAAAAGGCCAAAAATAAGAGTTTCCGAACCA 1380
Db |||||
QY 4785 CCAGCAATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGCTAAAGSTCAAGAGCTG 4844
Db |||||
QY 1381 CCCATCATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGCTGAAGTTTAAAGAGCTG 1440
Db |||||
QY 4845 ATGGTTCCCAAGTGTGTGAGGAGAAACCATCCATTTATGTCAGCAATTTGCAATAAGAA 4904
Db |||||
QY 1441 ATGGTTCCCAAGTGTGTGAGGAGAAAGCCAGCCATTTATGTTATGTCATTTGCAATAAGAG 1500
Db |||||
QY 4905 TATGATGACAAAGAAAGATTTTCTTCTTTC 4933
Db |||||
QY 1501 TATGATGACAAAGAAAGATTTTCTTCTTTC 1529
Db |||||

RESULT 4
U66672 LOCUS U66672 1346 bp mRNA linear HTC 23-JUL-2001
DEFINITION Homo sapiens clone EST90625 mRNA sequence.
ACCESSION U66672
VERSION U66672.1 GI:1906557
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1346)
AUTHORS Allikmets, R., Gerrard, B., Hutchinson, A. and Dean, M.
TITLE Characterization of the human ABC superfamily: isolation and
mapping of 21 new genes using the expressed sequence tags database
JOURNAL Hum. Mol. Genet. 5 (10), 1649-1655 (1996)
MEDLINE 97049974
PUBMED 8894702
REFERENCE 2 (bases 1 to 1346)
AUTHORS Allikmets, R., Gerrard, B. and Dean, M.
TITLE Direct Submission
SUBMITTED (12-AUG-1996) Human Genetics Section, National Cancer
Institute, NCI-FCRDC, Frederick, MD 21702, USA
JOURNAL
FEATURES
source
1..1346
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="EST90625"
/note="similar to ATP-binding cassette transporter"

BASE COUNT 451 a 206 c 290 g 395 t 4 others
ORIGIN

Query Match 18.1%; Score 1182; DB 11; Length 1346;
Best Local Similarity 96.5%; Pred. NO. 7.4e-164;
Matches 1306; Conservative 0; Mismatches 29; Indels 19; Gaps 9;

QY 4786 CAGCAATGAGGATGCAAGATGCAAGATGTCAAAGCTGAAAGCTTAAAGCTCAAGAGCTGA 4845
Db |||||
QY 1 CAGCAATGAGGATGCAAGATGCAAGATGTCAAAGCTGAAAGCTTAAAGCTCAAGAGCTGA 60
Db |||||
```

```
QY 4846 TGGGTGCCAGTGTGTGGAGAGAAACCAATCCATTATGTTGTCAGCAATTTGCATAAAGAAT 4905
Db 61 TGGGTGCCAGTGTGTGGAGAGAAACCAATCCATTATGTTGTCAGCAATTTGCATAAAGAAT 120
QY 4906 ATGATGACAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGGCAACTAAATACA 4965
Db 121 ATGATGACAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGGCAACTAAATACA 180
QY 4966 TCTCTTCTGTGTGCAAAAAAGGAGAGATCTTAGGACATATTGGGTCCAAATGGTCTGGCA 5025
Db 181 TCTCTTCTGTGTGCAAAAAAGGAGAGATCTTAGGACATATTGGGTCCAAATGGTCTGGCA 240
QY 5026 AAAGCAAAATTAATATTTCTGTGTGATATTGAACCAACTTCAGGCAGGATTTTT 5085
Db 241 AAAGCAAAATTAATATTTCTGTGTGATATTGAACCAACTTCAGGCAGGATTTTT 300
QY 5086 TAGGAGATTATCTTCAGAGCAAGTGAAGATGATGATTCACCTGAAGTGTATGGTTACT 5145
Db 301 TAGGAGATTATCTTCAGAGCAAGTGAAGATGATGATTCACCTGAAGTGTATGGTTACT 360
QY 5146 GTCTCTCAGATAAACCTTTGTGGCCAGATACATTCAGGAAACATTTTGAATTTATG 5205
Db 361 GTCTCTCAGATAAACCTTTGTGGCCAGATACATTCAGGAAACATTTTGAATTTATG 420
QY 5206 GAGCTGTCAAAGGAATGAGTGCAGTGCATGAAAGAGTCATAAGTCGAATACACATG 5265
Db 421 GAGCTGTCAAAGGAATGAGTGCAGTGCATGAAAGAGTCATAAGTCGAATACACATG 480
QY 5266 CACTTGATTTAAAGAACATCTTCAGAAAGCTGTAAGAAACTACCTGAGGATCAAC 5325
Db 481 CACTTGATTTAAAGAACATCTTCAGAAAGCTGTAAGAAACTACCTGAGGATCAAC 540
QY 5326 GAAAGTGTCTTGTCTAAGTATGCTAGGAACTCTCAGATTACTTCTCTAGATGAAC 5385
Db 541 CGAAAGTGTCTTGTCTAAGTATGCTAGGAACTCTCAGATTACTTCTCTAGATGAAC 600
QY 5386 CATCTACAGTATGATGCCAAA--GCCAAACAGCACATGTGGCAGCAATTCGAACTGC 5443
Db 601 CATCTACAGTATGATGCCAAA--GCCAAACAGCACATGTGGCAGCAATTCGAACTGC 660
QY 5444 ATTTAAACAGAAAGCGGTGCTATCTGACCACTCATCTATATGAGAGGAGGAGGC 5503
Db 661 AT-----NNAAGCGGGTCTGCTATCTGACCACTCACTATATGAGAGGAGGAGGC 711
QY 5504 TGCTGTGATCGAGTACATCATGTTGCTGGGAGTAAAGTATGATCGGAACAGTACA 5563
Db 712 TGCTGTGATCGAGTACATCATGTTGCTGGGAGTAAAGTATGATCGGAACAGTACA 771
QY 5564 ACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTAATTTGAA--GCACTGGA 5622
Db 772 ACATCTAAAGAGTAAATTTGGAAAG--NACTTTTGGAAATTAATTTGAA--GCACTGGA 829
QY 5623 TAGAAACCTTAGAAG--TAGACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCA 5681
Db 830 TAGAAACCTTAGAAGCTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCA 889
QY 5682 AGCGTGCAGAAAGCTTTTCTCTATTTTGGCTTATAAATCTCTAGGAGAGATGTTTCA 5741
Db 890 AGCGTGCAGAAAGCTTTTCTCTATTTTGGCTTATAAATCTCTAGGAGAGATGTTTCA 948
QY 5742 TCCCTTTTCAAACTTTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGGCAATTTGAAGA 5801
Db 949 TCCCTTTTCAAACTTTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGGCAATTTGAAGA 1007
QY 5802 TATAGCTTTTCTAAGCAATTTGGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAG 5861
Db 1008 TATAGC--TTTCTCAAGCAATTTGGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAG 1066
QY 5862 GAGGAGATATAGTGTGGAATTTTAAAGCAGCACTTTTGTGGGACCAACACAGAA 5921
Db 1067 GAGGAGATATAGTGTGGAATTTTAAAGCAGCACTTTTGTGGT--GGAAACCAACAGAA 1125
QY 5922 GATAGAGTAGTATTTTGAATTTTGTATTTGTTTGGTCTGCTTACTGGGACTTCTTTT 5981
```

```
Db 1126 GATAGAGTAGTATTTTGAATTTTGTATTTGCTGCTGCTTACTGGGACTTCTTTCTTTT 1185
QY 5982 CACTTAATTTTAACTTTTGGTTTAAAAGTTTATTTTGAATGTTAACTGGAACCAAG 6041
Db 1186 CACTTAATTTTAACTTTTGGTTTAAAAGTTTATTTTGAATGTTAACTGGAACCAAG 1245
QY 6042 AACGCACCTTGAATTTTCTTAAGCTCCTTAATTTGAATGCTGTTGTTGTTTCTT 6101
Db 1246 AACGCACCTTGAATTTTCTTAAGCTCCTTAATTTGAATGCTGTTGTTGTTTCTT 1305
QY 6102 TTCTTTAAATAAAACCTATGATATAATTAAGTGAA 6135
Db 1306 TTCTTTAAATAAAACCTATGATATAATTAAGTGAA 1339

RESULT 5
CD250953 850 bp mRNA linear EST 22-MAY-2003
LOCUS AGENCOURT 14214041 NIH MGC 179 Homo sapiens cDNA clone
DEFINITION IMAGE:30385589 5', mRNA sequence.
ACCESSION CD250953
VERSION CD250953.1 GI:31011419
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 850)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM452 row: 0 column: 06
High quality sequence stop: 719.
Location/Qualifiers
1..850
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30385589"
/tissue_type="Pituitary"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
/clone_lib="NIH MGC 179"
/notes="Organ: Brain; Vector: pCMV-SPORT6.1; Site 1: EcoRV
(destroyed); Site 2: NotI; Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon cloning
). Average insert size 1.1 Kb. Library was constructed by
(invitrogen). Note: this is a NIH MGC Library."
BASE COUNT 284 a 143 c 183 g 238 t 2 others
ORIGIN
Query Match 12.8%; Score 837; DB 14; Length 850;
Best Local Similarity 99.8%; Pred. No. 3.7e-113;
Matches 848; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 5014 ATGGTCTGCGAAAGACAAATTTATTAATA-TTCTGTTGGTGATATTGAACCAACTTCA 5072
Db 1 ATGGTCTGCGAAAGACAAATTTATTAANTTTCTGTTGGTGATATTGAACCAACTTCA 60
QY 5073 GGCCAGGTATTTTAGGAGATTATCTTCAGAGACAAGTGAAGATGATTCACCTGAAG 5132
Db 1125
```



```
QY 5510 TGATCGAGTAGCTATCATGTGCTCGGCAGGTTAAGATGATCGGAACAGTACAAACATCT 5569
Db 541 TGATCGAGTAGCTATCATGTGCTCGGCAGGTTAAGATGATCGGAACAGTACAAACATCT 600
QY 5570 AAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTAATTCGAAGGACTGGATAGAAA 5629
Db 601 AAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTAATTCGAAGGACTGGATAGAAA 660
QY 5630 CCTAGAAGTAGACCGGCTTCCAAAGAGAAATTCAGTATATTTTCCCAAATGCAAG-CCGTC 5688
Db 661 CCTAGAAGTAGACCGGCTTCCAAAGAGAAATTCAGTATATTTTCCCAAATGCAAGCCGTC 720
QY 5689 AGAAAGATTTTCTCTATTTTGGCTATATAAAT- CCTAAGGAAGATGTTCACT-CCCT 5746
Db 721 AGAAAGATTTTCTCTATTTTGGCTATATAAAT- CCTAAGGAAGATGTTCACT-CCCT 780
QY 5747 TTCACATCTTTTAA---GCTGGAGAGAGCTAAACAT-GCTTTTGGCAATTCAGAAAT 5802
Db 781 TTCACATCTTTTAA---GCTGGAGAGAGCTAAACATGCGCTTTTGGCAATTCAGAAAT 840
QY 5803 ATA 5805
Db 841 ATA 843

RESULT 7
BC029426 809 bp mRNA linear HTC 16-MAY-2002
LOCUS
DEFINITION Homo sapiens, Similar to ATP-binding cassette, sub-family A (ABC1), member 5, clone IMAGE:4723522, mRNA.
ACCESSION BC029426
VERSION BC029426.1 GI:20809520
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 809)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Direct Submission
JOURNAL Strausberg, R.
Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 41 Row: 1 Column: 5
This clone has the following problem: frame shifted.

FEATURES
source 1..809
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4723522"
/tissue_type="Liver"
/clone_lib="NIH MGC_76"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"

BASE COUNT 266 a 144 c 147 g 252 t
ORIGIN
```

```
Query Match 11.7%; Score 760.6; DB 11; Length 809;
Best Local Similarity 99.5%; Pred. No. 6.2e-102;
Matches 763; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 994 AGGTTTATTCAGAAAACATGTCACCTGCAATTTAGGAGGTTAGAGTTGGAGACAGACCA 1053
Db 38 AGGTTTATTCAGAAAACATGTCACCTGCAATTTAGGAGGTTAGAGTTGGAGACAGACCA 97
QY 1054 GAACACTTCTACTGAAGAATTAATTAATTAATGCAAGAACCAAAAGAGTAGTGTTCAGG 1113
Db 98 GAACACTTCTACTGAAGAATTAATTAATTAATGCAAGAACCAAAAGAGTAGTGTTCAGG 157
QY 1114 AAATTCCTTTTCCACTATTTTATTTTATTTGGTTAATATTAATTTAGCATGATCATCCAA 1173
Db 158 AAATTCCTTTTCCACTATTTTATTTTATTTGGTTAATATTAATTTAGCATGATCATCCAA 217
QY 1174 ATAAGAAATATGAAGAAGTCCCTAATATATAGAACTCAATCTATGGAAGATTTTACTCTTT 1233
Db 218 ATAAGAAATATGAAGAAGTCCCTAATATATAGAACTCAATCTATGGAAGATTTTACTCTTT 277
QY 1234 CTAATCTAATTTCTGGATATATCTCCAGTCACTAATATTAACAGCAGCATCATCAGAAAG 1293
Db 278 CTAATCTAATTTCTGGATATATCTCCAGTCACTAATATTAACAGCAGCATCATCAGAAAG 337
QY 1294 TGTCTACTCATCATCTACCTGATGTCTAATTAATTAATGCAAGAAATATACAAATGAAAAAGAAA 1353
Db 338 TGTCTACTCATCATCTACCTGATGTCTAATTAATTAATGCAAGAAATATACAAATGAAAAAGAAA 397
QY 1354 TGTTAATATCCAGTCTCTCTAAGCCGAGCAACTTTGTAGTGTGGTTTCAAAGACTCCA 1413
Db 398 TGTTAATATCCAGTCTCTCTAAGCCGAGCAACTTTGTAGTGTGGTTTCAAAGACTCCA 457
QY 1414 TGTCTCTATGAACCTTCTTTTCTTGATATGATTCAGTATCTCTATTTATATGATTT 1473
Db 458 TGTCTCTATGAACCTTCTTTTCTTGATATGATTCAGTATCTCTATTTATATGATTT 517
QY 1474 CAAGAGCTGGCTGTCTCAAAATCATGTGAGGCTGCTCAGTACTGCTCAGTGTTCACAG 1533
Db 518 CAAGAGCTGGCTGTCTCAAAATCATGTGAGGCTGCTCAGTACTGCTCAGTGTTCACAG 577
QY 1534 TTTTCAAGATCCATAGATGCTGCCATTAATACAGTTGAAGACCAATGTTTCTCTTTGGA 1593
Db 578 TTTTCAAGATCCATAGATGCTGCCATTAATACAGTTGAAGACCAATGTTTCTCTTTGGA 637
QY 1594 AGGAGCTGGAGTCAACTAAAGCTGTTTATTTGGGAGAACTGCTGTGTAGAAATAGATA 1653
Db 638 AGGAGCTGGAGTCAACTAAAGCTGTTTATTTATGGGAGAACTGCTGTGTAGAAATAGATA 697
QY 1654 CCTTTCCCGAGGAGTAATTTTAAATATACCTAGTTATAGCATTTTACCTTTTGGATAT 1713
Db 698 CCTTTCCCGAGGAGTAATTTTAAATATACCTAGTTATAGCATTTTACCTTTTGGATAT 757
QY 1714 TTTTGGCAATTCATATCGTAGCAGAAAAAGAAAAAATAAAGAA 1760
Db 758 TTTTGGCAATTCATATCGTAGCAGAAAAAGAAAAAATAAAGAA 804
```

```
RESULT 8
BQ441353
LOCUS
DEFINITION BQ441353 773 bp mRNA linear EST 24-MAY-2002
AGENCOURT_7838430 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6101304
5', mRNA sequence.
ACCESSION BQ441353
VERSION BQ441353.1 GI:21180429
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 773)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```



JOURNAL  
COMMENT

Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: rsb@nih.gov  
Tissue procurement: CLONTECH  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: AGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LCM2334 row: c column: 01  
High quality sequence stop: 537.  
Location/Qualifiers  
1..773  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6101304"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH MGC 82"  
/notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site\_1:  
SfiI (ggccattcgcc); Site\_2: SfiI (ggccattcgcc); 5' and  
3' adaptors were used in cloning as follows: 5' adaptor  
sequence: 5'-CAGCGCATTTATGCG-3' and 3' adaptor sequence:  
5'-ATTCTAGAGCCGCGGCACATG-dt(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size  
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained  
inserts by PCR. This library was enriched for full-length  
clones and was constructed by Clontech Laboratories (Palo  
Alto, CA)."

BASE COUNT 241 a 126 c 173 g 233 t  
ORIGINQuery Match 11.4%; Score 741.4; DB 13; Length 773;  
Best Local Similarity 98.4%; Pred. No. 4.1e-99;  
Matches 759; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 5323 AACGAAGTGTCTTTGCTCTAAGTATGCTAGGATCCTCAGATTACTTTCGTAGAG 5382  
DB 3 AAGAAAGTTGTGTTTGTCTAAGTATGCTAGGATCCTCAGATTACTTTCGTAGAG 62  
QY 5383 AACCATCTACAGGTATGGATCCCAAGCCAAACAGCACATGTGGCGAGCAATTCGAAC 5442  
DB 63 AACCATCTACAGGTATGGATCCCAAGCCAAACAGCACATGTGGCGAGCAATTCGAAC 122  
QY 5443 CATTTAAAAACAGAACGGCGGTCTATTCTGACCACTCTATATGAGGAGGAGG 5502  
DB 123 CATTTAAAAACAGAACGGCGGTCTATTCTGACCACTCTATATGAGGAGGAGG 182  
QY 5503 CTGCTGTGATCGAGTAGTATCATGTGTCTGGCAGTTAAGATGTCGGAACAGTAC 5562  
DB 183 CTGCTGTGATCGAGTAGTATCATGTGTCTGGCAGTTAAGATGTCGGAACAGTAC 242  
QY 5563 AACATCTAAAGAGTAAATTTGGAAAGCGTACTTTTGGAAATTAATTAAGGACTGGA 5622  
DB 243 AACATCTAAAGAGTAAATTTGGAAAGCGTACTTTTGGAAATTAATTAAGGACTGGA 302  
QY 5623 TAGAAAACTTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCAA 5682  
DB 303 TAGAAAACTTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCAA 362  
QY 5683 GCCGTACGGAAGTTTCTCTATTCTTTTGGCTTATATAATTCCTAAGGAGATGTTTCAGT 5742  
DB 363 GCCGTACGGAAGTTTCTCTATTCTTTTGGCTTATATAATTCCTAAGGAGATGTTTCAGT 422  
QY 5743 CCCTTTACCAATCTTTTAAAGCTGGAAGACTAAACATGCTTTTGCCATTCGAAGAT 5802  
DB 423 CCCTTTACCAATCTTTTAAAGCTGGAAGACTAAACATGCTTTTGCCATTCGAAGAT 482  
QY 5803 ATAGCTTTCTCAGCAACATTTGGAACAGGTTTTTGTAGAACTCCTCTAAAGAACAGAGG 5862  
DB 483 ATAGCTTTCTCAGCAACATTTGGAACAGGTTTTTGTAGAACTCCTCTAAAGAACAGAGG 542

QY 5863 AGGAAGATAATAGTTGTGGAACTTTAAACAGCACACACTTTGTGGAGAACACAAGAAG 5922  
DB 543 AGGAAGATAATAGTTGTGGAACTTTAAACAGCACACACTTTGTGGAGAACACAAGAAG 602  
QY 5923 ATAGAGTAGTATTTTGAATTTGTTATTTGTTTGGTCTGCTTACTGGGACTTCTTTTTC 5982  
DB 603 ATAGAGTAGTATTTTGAATTTGTTATTTGTTTGGGCTGCTTACTGGGACTTCTTTTTC 662  
QY 5983 ACTTAATTTTAACTTTGTTTAAAGAGTTTATTTGGAATGTTAACTGGAGAACCAAGA 6042  
DB 663 ACTTAATTTT-ACITTTGTTTAAAGAGTTTATTTGGAATGTTAACTGGAGAACCAAGA 721  
QY 6043 ACGCATTGAAATTTTCTAAGCTCTTAAATGAAATGCTGTGTTGTG 6093  
DB 722 ACGCATTGAAATTTTCTAAGCTCTTAAATGAAATGCTGGGTTGGG 772

RESULT 9  
BUG6259/cLOCUS 747 bp mRNA linear EST 07-OCT-2002  
DEFINITION UI-CF-DUI-adn-m-03-0-UI.s1 UI-CF-DUI Homo sapiens cDNA clone  
UI-CF-DUI-adn-m-03-0-UI 3', mRNA sequence.ACCESSION BUG6259  
VERSION BUG6259.1 GI:23540987  
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 747)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PubMed 8889548

COMMENT Contact: McCray, PB

McCray Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com) or from Open Biosystems

(www.openbiosystems.com).

Seq primer: M13 FORWARD

POLYA=yes.

Location/Qualifiers

1..747

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="UI-CF-DUI-adn-m-03-0-UI"

/tissue\_type="Primary Lung Epithelial Cells"

/dev\_stage="Adult"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone\_lib="UI-CF-DUI"

/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site\_1: EcoR I; Site\_2: Not I;

UI-CF-DUI is a normalized cDNA library containing the

following tissue(s): Primary Lung Epithelial Cells The

library was constructed according to Bonaldo, Lennon and

Soares, Genome Research, 6:791-806, 1996. First strand

cDNA synthesis was primed with an oligo-dT primer

containing a Not I site. Double stranded cDNA was ligated

to an EcoR I adaptor, digested with Not I, and cloned

directionally into pT7T3-Pac vector. The oligonucleotide

used to prime the synthesis of first-strand cDNA contains

a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is GCGTGTAGGC.

TAG LIB=UI-CF-DUI

TAG\_TISSUE=Lung Epithelial Cells Tissue nos 359-368

TAG\_SEQ=GCGTGTAGGC

BASE COUNT 230 a 153 c 112 g 252 t

ORIGIN

Query Match 11.2%; Score 731; DB 13; Length 747;  
Best Local Similarity 100.0%; Pred. No. 1.4e-97;  
Matches 731; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5405 CAAAGCCAAACAGCACATGTGGCGAGCAATTCGAATCGCATTTTAAACACAGAAAGCGGC 5464

Db 747 CAAAGCCAAACAGCACATGTGGCGAGCAATTCGAATCGCATTTTAAACACAGAAAGCGGC 688

QY 5465 TGCTATTCGACACTCATTATGAGGAGGAGCGCTGTCTGTGATCGAGTAGCTAT 5524

Db 687 TGCTATTCGACACTCATTATGAGGAGGAGCGCTGTCTGTGATCGAGTAGCTAT 628

QY 5525 CATGGTGTCTGGCAGTTAAGATGTATCGGAACAGTACACATCTAAAGAGTAAATTGG 5584

Db 627 CATGGTGTCTGGCAGTTAAGATGTATCGGAACAGTACACATCTAAAGAGTAAATTGG 568

QY 5585 AAAAGGCTACTTTTGGAAATTAATTTGAAGGACTGGATAGAAAACCTAGAACTAGACCG 5644

Db 567 AAAAGGCTACTTTTGGAAATTAATTTGAAGGACTGGATAGAAAACCTAGAACTAGACCG 508

QY 5645 CTTTCAAAGAAATTCAGTATATTTTCCAAATGCAAGCGCTCAGGAAAGTTTCTTC 5704

Db 507 CTTTCAAAGAAATTCAGTATATTTTCCAAATGCAAGCGCTCAGGAAAGTTTCTTC 448

QY 5705 TATTTGGCTTATAAATTTCTAAGGAGATGTTTCAGTCCCTTCACAATCTTTTAA 5764

Db 447 TATTTGGCTTATAAATTTCTAAGGAGATGTTTCAGTCCCTTCACAATCTTTTAA 388

QY 5765 GCTGGAAGAGCTAAACATGCTTTTGCAATTTGAAGATATAGCTTTTCTCAAGCAACAT 5824

Db 387 GCTGGAAGAGCTAAACATGCTTTTGCAATTTGAAGATATAGCTTTTCTCAAGCAACAT 328

QY 5825 GGAACAGGTTTTGTAGAACTCACTAAAGAACAGAGGAGGAAGATAATAGTTGTGAAC 5884

Db 327 GGAACAGGTTTTGTAGAACTCACTAAAGAACAGAGGAGGAAGATAATAGTTGTGAAC 268

QY 5885 TTTAAACAGCACATTTGGTGGGACGAAACACAGAGATAGAGTAGTATTTGAATTG 5944

Db 267 TTTAAACAGCACATTTGGTGGGACGAAACACAGAGATAGAGTAGTATTTGAATTG 208

QY 5945 TATTGTTCGGTCTGCTTACTTGGGACTTCTTTTCTTTTCACTTAATTTTAACCTTTGGTTA 6004

Db 207 TATTGTTCGGTCTGCTTACTTGGGACTTCTTTTCTTTTCACTTAATTTTAACCTTTGGTTA 148

QY 6005 AAAAGTTTTTTTATTTGAATGGTAACTGAGAACCAAGACCGCACTTGAATTTTCTAAG 6064

Db 147 AAAAGTTTTTTTATTTGAATGGTAACTGAGAACCAAGACCGCACTTGAATTTTCTAAG 88

QY 6065 CTCTTAATTTGAATGCTGTGTTGTGTTGTTGTTTCTTTCTTTTAAATAAAGTAGTAT 6124

Db 87 CTCTTAATTTGAATGCTGTGTTGTGTTGTTTCTTTCTTTTAAATAAAGTAGTAT 28

QY 6125 AATTAAGTGAA 6135

Db 27 AATTAAGTGAA 17

RESULT 10  
BG564439  
LOCUS  
DEFINITION 602584385F1 NIH\_MGC\_76 Homo sapiens cdna clone IMAGE:4712157 5',  
mRNA sequence.  
ACCESSION BG564439  
VERSION BG564439.1 GI:13572091

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EST.

Homo sapiens (human)

Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 771)

NIH-MGC http://mhc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: CLONTECH Laboratories, Inc.

CDNA Library Preparation: CLONTECH Laboratories, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCMI553 row: i column: 22

High quality sequence stop: 763.

FEATURES

source

1..771

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NIH\_MGC\_76"

/notes="Organ: liver; Vector: pDNR-LIB (Clontech); Site: 1:

SfiI (ggcgctcgcc); Site 2: SfiI (ggcgctcgcc); 5' and

3' adaptors were used in cloning as follows: 5' adaptor

sequence: 5'-CACGCCATTATGCGC-3' and 3' adaptor sequence:

5'-ATTCTAGAGCGGCGCGCATG-dt(30)BN-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size 1.85

kb (range 1.0-4.0 kb). 15/15 colonies contained inserts

by PCR. This library was enriched for full-length clones

and was constructed by Clontech Laboratories (Palo Alto,

CA). Note: this is a NIH MGC Library."

BASE COUNT 275 a 130 c 167 g 199 t

ORIGIN

Query Match 11.1%; Score 722.4; DB 10; Length 771;

Best Local Similarity 99.4%; Pred. No. 2.5e-96;

Matches 767; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 4721 AAGAAAAGATCCCTTTTTCAGAAACCTTTCAACGAAGTCTTAAATAAGGAGCTCCAGA 4780

Db 1 AAGAAAAGATCCCTTTTTCAGAAACCTTTCAACGAAGTCTTAAATAAGGAGCTCCAGA 60

QY 4781 ACCACACAGACAATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGACTAAAGGTCAAAGA 4840

Db 61 ACCACACAGACAATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGACTAAAGGTCAAAGA 120

QY 4841 GCTGATGGGTGTCAGTGTGTGAGAGAAACCATCATTTATGCTCAGCAATTTGCATAA 4900

Db 121 GCTGATGGGTGTCAGTGTGTGAGAGAAACCATCATTTATGCTCAGCAATTTGCATAA 180

QY 4901 AGAATATGATGACAAGAAAGATTTTCTTCTTCAAGAAAAGTAAAGAAAGTGCACATAA 4960

Db 181 AGAATATGATGACAAGAAAGATTTTCTTCTTCAAGAAAAGTAAAGAAAGTGCACATAA 240

QY 4961 ATACATCTCTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGGTCAAATGTTGC 5020

Db 241 ATACATCTCTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGGTCAAATGTTGC 300

QY 5021 TGGCAAAAGCACAAATTAATTAATTTCTGGTGTGATATTGAACCAACTTCAGGCCAGGT 5080

Db 301 TGGCAAAAGCACAAATTAATTAATTTCTGGTGTGATATTGAACCAACTTCAGGCCAGGT 360

QY 5081 ATTTTGTAGGAGATATTCTTTTCAGAGACAAGTGAAGATGATGATTTCTGTAAGTGTATGGG 5140

Db 361 ATTTTGTAGGAGATATTCTTTTCAGAGACAAGTGAAGATGATGATTTCTGTAAGTGTATGGG 420



```

VERSION BG435656.1 GI:13342162
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 756)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
          Email: gchapb@mail.nih.gov
          Tissue Procurement: CLONTECH Laboratories, Inc.
          cDNA Library Preparation: CLONTECH Laboratories, Inc.
          DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.lnlni.gov
          Plate: LCM1339 row: c column: 12
          High quality sequence stop: 706.

FEATURES             source
   Location/Qualifiers
     1..756
       /organism="Homo sapiens"
       /mol_type="mRNA"
       /db_xref="taxon:9606"
       /clone="IMAGE:4604099"
       /lab_host="DH10B (TI phage-resistant)"
       /clone_lib="NIH_MGC_79"
       /notes="Organ: Placenta; Vector: pDNR-LIB (Clontech);
       Site: SfiI (ggccgctcgcc); Site 2: SfiI (ggccgctatggcc
       ); 5' and 3' adaptors were used in cloning as follows: 5'
       adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor
       sequence: 5'-ATTCGAGGCGCGAGCGCCGACATG-dT(30)BN-3'
       (where B = A, C, or G and N = A, C, G, or T). Average
       insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies
       contained inserts by PCR. This library was enriched for
       full-length clones and was constructed by Clontech
       Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
       Library."
BASE COUNT 269 a 123 c 168 g 196 t

Query Match 10.4%; Score 681.6; DB 10; Length 756;
Best Local Similarity 97.1%; Pred. No. 2.5e-90;
Matches 726; Conservative 0; Mismatches 19; Indels 3; Gaps 3;

QY 5187 GAACATTTTGAATTTATGAGCTGTCAAAGGAATGATGTCACATGACATGAAAGAGTC 5246
    |||
    |||
    |||
Db 1 GAACATTTTGAATTTATGAGCTGTCAAAGGAATGATGTCACATGACATGAAAGAGTC 60

QY 5247 ATAAGTCGAATACACATGCACCTGATTTAAAGACATCTTCAGAGACTGTAAAGAA 5306
    |||
    |||
    |||
Db 61 ATAAGTCGAATACACATGCACCTGATTTAAAGACATCTTCAGAGACTGTAAAGAA 120

QY 5307 CTACCTGCAGGAATCAAAAGAAAGTTGTGTTTGTCTTAAGTATGCTAGGAACTCTCAG 5366
    |||
    |||
    |||
Db 121 CTACCTGCAGGAATCAAAAGAAAGTTGTGTTTGTCTTAAGTATGCTAGGAACTCTCAG 180

QY 5367 ATTACTTTGTAGATGAACCATCTACAGGTATGGATCCCAAGCCAAACAGCAGCATGTGG 5426
    |||
    |||
    |||
Db 181 ATTACTTTGTAGATGAACCATCTACAGGTATGGATCCCAAGCCAAACAGCAGCATGTGG 240

QY 5427 CGAGCAATTCGAATCAATTTAAACAGAGAGGGGCTGCTATTCTGACCACTCACTAT 5486
    |||
    |||
    |||
Db 241 CGAGCAATTCGAATCAATTTAAACAGAGAGGGGCTGCTATTCTGACCACTCACTAT 300

QY 5487 ATGAGGAGGAGCAGAGGCTGTCTGTGATCGAGTAGCTATCATGTGTCTGGGCACTTAAGA 5546
    |||
    |||
    |||
Db 301 ATGAGGAGGAGCAGAGGCTGTCTGTGATCGAGTAGCTATCATGTGTCTGGGCACTTAAGA 360

QY 5547 TGTATCGGAACAGTACAACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAAAT 5606
    |||
    |||
    |||

```

```

Db 361 TGTATCGGAACAGTACAACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAAAT 420.
QY 5607 AAATTGAAGGACTGGATAGAAAACCTAGAGTAGACCGCTTCAAAGAGAGAAATTCAGTAT 5666
    |||
    |||
    |||
Db 421 AAATTGAAGGACTGGATAGAAAACCTAGAGTAGACCGCTTCAAAGAGAGAAATTCAGTAT 480
    |||
    |||
    |||
QY 5667 ATTTTCCCAATGCAAGCCGCTCAGGAA-AGTTTTTCTTCTATTTTGGCTTATAAAATTC 5725
    |||
    |||
    |||
Db 481 ATTTTCCCAATGCAAGCCGCTCAGGAAACAGTTTTTCTTCTATTTGGCTTATAAAATTC 540
    |||
    |||
    |||
QY 5726 TAAGGAAGATGTTCACTCCCTTTCACAAATCTTTTTTAAGCTGGAAGAAGCT-AAACATG 5784
    |||
    |||
    |||
Db 541 TAAGGAAGATGTTCACTCCCTTTCACAAATCTTTTTTAAGCTGGAAGAAGCTAAACATG 600
    |||
    |||
    |||
QY 5785 CTTTTCGCCATTGAAGAATATAGCTTTTCTCAAGCAACATTTGGAAAGGTTTTTGTAGAAC 5844
    |||
    |||
    |||
Db 601 CTTTTCGCCATTGAAGAATATAGCTTTTCTCAAGCAACATTTGGAAAGGTTTTTGTAGAAC 660
    |||
    |||
    |||
QY 5845 TCAC-TAAAGAAACAAGAGGAGGAGGATAATAGTTGTGGAACTTTAAACAGCACACTTTGG 5903
    |||
    |||
    |||
Db 661 TCAC-TAAAGAAACAAGAGGAGGAGGATAATAGTTGTGGAACTTTAAACAGCACACTTTGG 720
    |||
    |||
    |||
QY 5904 TGGGAACGACACACAAGATAGAGTAG 5931
    |||
    |||
    |||
Db 721 GGGAAACGACACACAAGATAGAGTAG 748
    |||
    |||
    |||

RESULT 13
AL700140
LOCUS
DEFINITION DKF2p686H03117_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION AL700140
VERSION AL700140.1 GI:19620673
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 744)
          Wiemann,S.
          OTTENWAEIDER,B., Obermaier,B., Mewes,H.W., Weill,B. and Wiemann
          .S.)
          Unpublished
          Contact: Ottenwaeider B
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MedGenomix (Martinsried/Germany) within the CDNA
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKF2p686H03117) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             source
   Location/Qualifiers
     1..744
       /organism="Homo sapiens"
       /mol_type="mRNA"
       /db_xref="taxon:9606"
       /clone="DKF2p686H03117"
       /tissue_type="human skeletal muscle"
       /dev_stage="adult"
       /lab_host="DH10B"
       /clone_lib="686 (synonym: hlcc3)"
       /note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
       cDNA-collection"
BASE COUNT 247 a 94 c 138 g 265 t

Query Match 10.4%; Score 680.4; DB 9; Length 744;

```

Best Local Similarity 99.2%; Pred. No. 3.8e-90;		Matches 736; Conservative 0; Mismatches 1; Indels 5; Gaps 5;	
QY	5789	TGCCATTGAAGATATAGCTTTCTCAAGCAACATTGGACAGCTTTTGTAGAACTCAC	5848
Db	1	TGCCATTGAAGATATAGCTTTCTCAAGCAACATTGGACAGCTTTTGTAGAACTCAC	60
QY	5849	TAAAGAACAGAGGAGGAGATATAGTGTGCACTTTAAACAGACACACTTTGGTGGGA	5908
Db	61	TAAAGAACAGAGGAGGAGATATAGTGTGCACTTTAAACAGACACACTTTGGTGGGA	120
QY	5909	ACGAACACAGAGATAGAGTAGTATTTGAAATTTGTTGTTGCTGCTTACTGGGA	5968
Db	121	ACGAACACAGAGATAGAGTAGTATTTGAAATTTGTTGTTGCTGCTTACTGGGA	180
QY	5969	CTTCTTTCTTTTCACTTAATTTAACTTTGGTTTAAAGTTTATTTTGAATCGTAA	6028
Db	181	CTTCTTTCTTTTCACTTAATTTAACTTTGGTTTAAAGTTTATTTTGAATCGTAA	240
QY	6029	CTGGAGAACCAAGACGCACTTGAATTTTCTAAGCTCTTAATTTGAATGCTGTGTT	6088
Db	241	CTGGAGAACCAAGACGCACTTGAATTTTCTAAGCTCTTAATTTGAATGCTGTGTT	300
QY	6089	GTGTGTTTCTTTCTTTTAAATAAAGCTATGTATATAATTAAGTGAAGCTGCTGTTGT	6148
Db	301	GTGTGTTTCTTTCTTTTAAATAAAGCTATGTATATAATTAAGTGAAGCTGCTGTTGT	360
QY	6149	ATTGAAGATATATTGAATATATAGTTTGTATGTCATCTTTTCCACCATTCAGAACAGTG	6208
Db	361	ATTGAAGATATATTGAATATATAGTTTGTATGTCATCTTTTCCACCATTCAGAACAGTG	420
QY	6209	CTTCTGAATTTGATTTAAAGCAATTTGAATAGATAGTTTAA-TTTTAACTTATCTT	6267
Db	421	CTTCTGAATTTGATTTAAAGCAATTTGAATAGATAGTTTAA-TTTTAACTTATCTT	480
QY	6268	TAAAGTTTATGCCATCTTCTTAAATAAGTACGTAATGTTTCAATCTTAATAAATAAATAAT	6327
Db	481	TAAAGTTTATGCCATCTTCTTAAATAAGTACGTAATGTTTCAATCTTAATAAATAAATAAT	540
QY	6328	ACATACTAATGATAGAAAGATACATAAAGCAAT-GTGAAGTTTCTTGTCTTCTTCTT	6386
Db	541	TCATACTAATGATAGAAAGATACATAAAGCAATGTTGAAGTTTCTTGTCTTCTTCTT	600
QY	6387	TTTAAATTTCTAAAGACCACTTTGAT-GGAAGTTCTCATCGTAAAGCTGAAGTGA	6445
Db	601	TTTAAATTTCTAAAGACCACTTTGATGGAAGTTCTCATCGTAAAGCTGAAGTGA	660
QY	6446	AGCACTAGGAAATCTCAATATAGAGATT-GAGGAAAGTTTATATCCACTA-GGTGGCAGT	6503
Db	661	AGCACTAGGAAATCTCAATATAGAGATTGAGGAAAGTTTATATCCACTAGGTGGCAGT	720
QY	6504	CATTGATCATATAAGTGAAT 6525	
Db	721	CATTGATCATATAAGTGAAT 742	
RESULT 14			
BQ774317/c			
LOCUS			
DEFINITION			
UI-H-EZ1-bca-d-17-0-UI.s1 NCI CGAP Ch2 Homo sapiens cDNA clone			
UI-H-EZ1-bca-d-17-0-UI 3', mRNA sequence.			
ACCESSION			
BQ774317			
KEYWORDS			
EST.			
SOURCE			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 697)			
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.			
Tumor Gene Index			
JOURNAL			
Unpublished			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-remail.nih.gov			
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbiterian, Dept. of			
Orthopaedics			
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa			
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa			
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa			
Clone Distribution: Clone distribution information can be obtained			
from Dr. M. Bento Soares, bento-soares@uiowa.edu			
The following repetitive elements were found in this cDNA			
sequence: 1-25, >AT-rich#low_complexity (matched complement)			
Seq primer: M13 FORWARD			
POLYA=Yes.			
FEATURES			
source			
1. 697			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="UI-H-EZ1-bca-d-17-0-UI"			
/tissue_type="Chondrosarcoma Grade II"			
/dev_stage="Adult"			
/lab_host="DH10B (Life Technologies)"			
/clone_lib="NCI CGAP Ch2"			
/note="Organ: Left Pelvis; Vector: pTT73-Pac (Pharmacia)			
with a modified polylinker; Site 1: Ecor I; Site 2: Not I;			
NCI CGAP Ch2 is a normalized cDNA library containing the			
following tissue(s): Chondrosarcoma Grade II. The library			
was constructed according to Bonaldo, Lennon and Soares,			
Genome Research, 6:791-806, 1996. First strand cDNA			
synthesis was primed with an oligo-dT primer containing a			
Not I site. Double stranded cDNA was ligated to an Ecor I			
adaptor, digested with Not I, and cloned directionally			
into pTT73-Pac vector. The oligonucleotide used to prime			
the synthesis of first-strand cDNA contains a library tag			
sequence that is located between the Not I site and the			
(dT)18 tail. The sequence tag for this library is			
TGATCAGCT.			
TAG LIB=UI-H-EZ1			
TAG_TISSUE=grade-2-chondrosarcoma			
TAG_SEQ=ATCTAATG			
BASE COUNT 220 a 143 c 101 g 233 t			
ORIGIN			
Query Match 10.4%; Score 676.4; DB 13; Length 697;			
Best Local Similarity 99.1%; Pred. No. 1.5e-89;			
Matches 680; Conservative 0; Mismatches 6; Indels 0; Gaps 0;			
QY	5450	AAACAGAAAGCGGCTGCTATTCTGACCCTCATATATGGAGGAGGAGGCTGTCTG	5509
Db	697	AAACAGAAAGCGGCTGCTATTCTGACCCTCATATATGGAGGAGGAGGCTGTCTG	638
QY	5510	TGATCAGTAGCTATCTGGGCTGCTGGGAGTATAGATGTCGGAACAGTACACATCT	5569
Db	637	TGATCAGTAGCTATCTGGGCTGCTGGGAGTATAGATGTCGGAACAGTACACATCT	578
QY	5570	AAAGAGTAAATTGGAAAGGCTACTTTTGAAATTAATTAAGGAGTATAGAGAA	5629
Db	577	AAAGAGTAAATTGGAAAGGCTACTTTTGAAATTAATTAAGGAGTATAGAGAA	518
QY	5630	CCTAGAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCAAATGCAAGCGTCA	5689
Db	517	CCTAGAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCAAATGCAAGCGTCA	458
QY	5690	GGAAAGTTTCTTCTATTTGGCTTATAAATTCCTAAGGAGATGTTTCAGTCCCTTTC	5749
Db	457	GGAAAGTTTCTTCTATTTGGCTTATAAATTCCTAAGGAGATGTTTCAGTCCCTTTC	398
QY	5750	ACAACTCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGCCATTGAAGATATAGCTT	5809
Db	397	ACAACTCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGCCATTGAAGATATAGCTT	338
QY	5810	TTCTCAAGCAACATTTGGAAAGGTTTTTGTAGAACTCACTAAAGAACAGAGGAGAGA	5869



**THIS PAGE BLANK (USPTO)**



Pred. No. is the number of results predicted by chance to have a

RESULT	1
AX473847	
LOCUS	
DEFINITION	Sequence 1 from Patent WO0246458.
ACCESSION	AX473847
VERSION	AX473847.1 GI:22208006
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1
AUTHORS	Denefle,P., Rosier-Montus,M.F., Prades,C., Arnould-Reguinne,I., Duverger,N., Allikmets,R. and bean,M. Nucleic acids of the human abca5, abca9, and abcal0 genes,
TITLE	

vectors containing such nucleic acids and uses thereof

JOURNAL Patent: WO 0246458-A.1 13-JUN-2002;  
Aventis Pharma S.A. (FR); The Secretary, Department of Health and  
Human Services (US)

FEATURES  
source Location/Qualifiers

1.6525  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 1973 a 1108 c 1258 g 2185 t 1 others

ORIGIN

Query Match 100.0%; Score 6524; DB 6; Length 6525;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 6525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAATGTTGATATTTCTCTTAGCAGCTGTCTAACAGGTTAGGTTCAAGTCTAAGTTT 60

Db 1 AAAATGTTGATATTTCTCTTAGCAGCTGTCTAACAGGTTAGGTTCAAGTCTAAGTTT 60

Qy 61 CTACCCACATCTTTGACCTGTAGTGTCTATTTAGTATTTTCAAAAACCTTTTGCAG 120

Db 61 CTACCCACATCTTTGACCTGTAGTGTCTATTTAGTATTTTCAAAAACCTTTTGCAG 120

Qy 121 TACCTTTTGGTCTGTCTGTGTGCTTGCAGTGAACAGTCTGGATTGGACAGTGGT 180

Db 121 TACCTTTTGGTCTGTCTGTGTGCTTGCAGTGAACAGTCTGGATTGGACAGTGGT 180

Qy 181 CTGTCTGTAGTTAGTTCTCAAGCTTTGTGCACACTAATAGGATGGATTATGTATG 240

Db 181 CTGTCTGTAGTTAGTTCTCAAGCTTTGTGCACACTAATAGGATGGATTATGTATG 240

Qy 241 TCAGCTTGGGAATTAACAGAAATTAACAACTTTTACAGCTCTTCTGTAGCTC 300

Db 241 TCAGCTTGGGAATTAACAGAAATTAACAACTTTTACAGCTCTTCTGTAGCTC 300

Qy 301 TCTTTCTATTTGTTCCCTCTTCTACTTTTGTCTGCTGCTGCTGCTTCTATCCTCC 360

Db 301 TCTTTCTATTTGTTCCCTCTTCTACTTTTGTCTGCTGCTGCTGCTTCTATCCTCC 360

Qy 361 AGCCAGAGCTAGTGTATTTCTCCATTTGTGTATACACTTTGTGCAGCTGCAACCCAC 420

Db 361 AGCCAGAGCTAGTGTATTTCTCCATTTGTGTATACACTTTGTGCAGCTGCAACCCAC 420

Qy 421 CATATCCAGGCCCAATGTAGAGGTAGAGAGAAAGCAAGGATTTGGCTCATCC 480

Db 421 CATATCCAGGCCCAATGTAGAGGTAGAGAGAAAGCAAGGATTTGGCTCATCC 480

Qy 481 TCTTACACGATAGTTCATTTGATAGAGAAAGGTTTTCTGCTCAGAGTGTGGCT 540

Db 481 TCTTACACGATAGTTCATTTGATAGAGAAAGGTTTTCTGCTCAGAGTGTGGCT 540

Qy 541 GCCTAGGCTTTTGTACTGTAGTCTGGCCCTGTACCAAGGATTTGCTTGCATGTGGGG 600

Db 541 GCCTAGGCTTTTGTACTGTAGTCTGGCCCTGTACCAAGGATTTGCTTGCATGTGGGG 600

Qy 601 ATACAGAGAAATTCAGAAAGAAAAGATTTGCTATTTCTACATTTCCCTGAGCAT 660

Db 601 ATACAGAGAAATTCAGAAAGAAAAGATTTGCTATTTCTACATTTCCCTGAGCAT 660

Qy 661 AAGACCTTCCCTGCTCCATTTCAATTTCAAGCTTAAGGCTTTCTTGGAGCTGCTCTGT 720

Db 661 AAGACCTTCCCTGCTCCATTTCAATTTCAAGCTTAAGGCTTTCTTGGAGCTGCTCTGT 720

Qy 721 GGGCGGTTGGGAGATACCAAGAGAGAAAAGTACCACTGTTGATATGGTATTTCAA 780

Db 721 GGGCGGTTGGGAGATACCAAGAGAGAAAAGTACCACTGTTGATATGGTATTTCAA 780

Qy 781 ATTCTGTCTACCTATTTACATGCTGTTTACTTTTACAGAGCTGACAGATTTGCTCT 840

Db 781 ATTCTGTCTACCTATTTACATGCTGTTTACTTTTACAGAGCTGACAGATTTGCTCT 840

Qy 841 CCATGCATTTCTGCTCCAGTTTCTTAAGAGACAGCTTGGAGTATGCTTAATCCATCTTAC 900

841 CCATGCATTTCTGCTCCAGTTTCTTAAGAGACAGCTTGGAGTATGCTTAATCCATCTTAC 900

901 CTGGGACTGAACAGCTGCTTTATTTTGCCTTAAATAATACATCAGTTTACTGCTGGC 960

901 CTGGGACTGAACAGCTGCTTTATTTTGCCTTAAATAATACATCAGTTTACTGCTGGC 960

961 TCCGGGTTTGTGTTGTTTCTCTTTAATAGGTTTATTCAGAAAAACATGTCACCTG 1020

961 TCCGGGTTTGTGTTGTTTCTCTTTAATAGGTTTATTCAGAAAAACATGTCACCTG 1020

1021 CAATPAGGAGGTAGGAGTGGAGACAGACAGAAACACTTCTCTAAGAAATTAATCTTAA 1080

1021 CAATPAGGAGGTAGGAGTGGAGACAGACAGAAACACTTCTCTAAGAAATTAATCTTAA 1080

1081 TTAATGACAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACATATTTTTTTAT 1140

1081 TTAATGACAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACATATTTTTTTAT 1140

1141 TTTGGTTAATTAATAGCATGATGCATCAAAATAAGAAATTAAGAAAGTGCCTAATA 1200

1141 TTTGGTTAATTAATAGCATGATGCATCAAAATAAGAAATTAAGAAAGTGCCTAATA 1200

1201 TAGAACTCAATCTTATGACCAAGTTTACTCTTTCTAATCTAATCTTGGATATATCTCCAG 1260

1201 TAGAACTCAATCTTATGACCAAGTTTACTCTTTCTAATCTAATCTTGGATATATCTCCAG 1260

1261 TGACTTAATTAACAAGCAGCATCATGCAAGAAAGTGTCTACTGATCATCTACTGATGCA 1320

1261 TGACTTAATTAACAAGCAGCATCATGCAAGAAAGTGTCTACTGATCATCTACTGATGCA 1320

1321 TAAATPACTGAAGATATACAAATGAAGAAAGAAATGTTAAACATCCAGTCTCTTAAGCCGA 1380

1321 TAAATPACTGAAGATATACAAATGAAGAAAGAAATGTTAAACATCCAGTCTCTTAAGCCGA 1380

1381 GCAACTTTGTAGGTGTGGTTTCAAGACTCCATGTCCTATGAACCTCGTTTCTCTG 1440

1381 GCAACTTTGTAGGTGTGGTTTCAAGACTCCATGTCCTATGAACCTCGTTTCTCTG 1440

1441 ATATGATTTCCAGTATCTCTATTTATATGAAAGTCTCAAGAGCTGGCTGTTCAAAATCATGTG 1500

1441 ATATGATTTCCAGTATCTCTATTTATATGAAAGTCTCAAGAGCTGGCTGTTCAAAATCATGTG 1500

1501 AGGCTGCTCAGTACTGCTCCTCAGGTTTCAAGCTTTCAAGAGCTGGCTGTTCAAAATCATGTG 1560

1501 AGGCTGCTCAGTACTGCTCCTCAGGTTTCAAGCTTTCAAGAGCTGGCTGTTCAAAATCATGTG 1560

1561 TTATACAGTTGAAGACCAATGTTCTCTTTGGAAGAGCTGGAGTCAACTAAAGCTGTTA 1620

1561 TTATACAGTTGAAGACCAATGTTCTCTTTGGAAGAGCTGGAGTCAACTAAAGCTGTTA 1620

1621 TTATGGGAGAACTGCTGTTGTAAGAAATAGATACCTTTCCCGAGGAGTAAATTAATAT 1680

1621 TTATGGGAGAACTGCTGTTGTAAGAAATAGATACCTTTCCCGAGGAGTAAATTAATAT 1680

1681 ACCTAGTTATAGCATTTTCACTTTTGGATATCTTTTGGCAATTCATATCGTAGCAGAAA 1740

1681 ACCTAGTTATAGCATTTTCACTTTTGGATATCTTTTGGCAATTCATATCGTAGCAGAAA 1740

1741 AAGAAAAAATAAAGAAATTTTAAAGATAAATGGGACTTTCATGATATCTGCTTTTGGC 1800

1741 AAGAAAAAATAAAGAAATTTTAAAGATAAATGGGACTTTCATGATATCTGCTTTTGGC 1800

1801 TTTCTGCTGCTTCTATATACAGTTTAAATTTTCTTATGCTCCCTTCTATGAGCAGTCA 1860

1801 TTTCTGCTGCTTCTATATACAGTTTAAATTTTCTTATGCTCCCTTCTATGAGCAGTCA 1860

1861 TTGCGACAGCTTCTTTGTTATTTCTCTCAAGTAGCAGCATTTGTGATATTTCTGCTTTTTT 1920

1861 TTGCGACAGCTTCTTTGTTATTTCTCTCAAGTAGCAGCATTTGTGATATTTCTGCTTTTTT 1920

1921 TCCTTTATGGATTAATCATCTGTAATTTTGTCTTAATGCTGACACCTCTTTTAAAAAAT 1980

1921 TCCTTTATGGATTAATCATCTGTAATTTTGTCTTAATGCTGACACCTCTTTTAAAAAAT 1980

Db 1921 TCCTTTATGGATTATCATCTGTATTTTTTTCCTTTAAATGCTGACACCTCTTTTTAAAAAAT 1980  
Qy 1981 CAAAAATGTCGGAATAGTTTGAATTTTTTCTGTACTGTGGCTTTTGGATTTATTTGGCCTTA 2040  
Db 1981 CAAAAATGTCGGAATAGTTTGAATTTTTTCTGTACTGTGGCTTTTGGATTTATTTGGCCTTA 2040  
Qy 2041 TGATAATCCTCATAGAAAGTTTTCCCAAATCGTTAGTGTGGCTTTTCAGTCCCTTTCTGTC 2100  
Db 2041 TGATAATCCTCATAGAAAGTTTTCCCAAATCGTTAGTGTGGCTTTTCAGTCCCTTTCTGTC 2100  
Qy 2101 ACTGTACTTTTGTGATTTGATTTGTCACAGCTCATGCAATTTTAGAGATTTTAAATGAAGGTG 2160  
Db 2101 ACTGTACTTTTGTGATTTGTCACAGCTCATGCAATTTTAGAGATTTTAAATGAAGGTG 2160  
Qy 2161 CTTCAATTTTCAAAATTTGACTGTCAGGCCCATATCCTCTAAATATTACAATTTATCATGCTCA 2220  
Db 2161 CTTCAATTTTCAAAATTTGACTGTCAGGCCCATATCCTCTAAATATTACAATTTATCATGCTCA 2220  
Qy 2221 CACTTAATAGTATATCTATGCTCTTGTGCTGTCTATCTTGATCAAGTCAATTCAGGGG 2280  
Db 2221 CACTTAATAGTATATCTATGCTCTTGTGCTGTCTATCTTGATCAAGTCAATTCAGGGG 2280  
Qy 2281 AATTGGCTTTACCGAGATCATCTTTATATTTCTGAGCCCTTCATATTTGGTCAAGAGTA 2340  
Db 2281 AATTGGCTTTACCGAGATCATCTTTATATTTCTGAGCCCTTCATATTTGGTCAAGAGTA 2340  
Qy 2341 AAAGAAATTAATGAGGAGTATCAGAGGCCAATGTTAATGGAATATTAGTTTGTAGTGAAA 2400  
Db 2341 AAAGAAATTAATGAGGAGTATCAGAGGCCAATGTTAATGGAATATTAGTTTGTAGTGAAA 2400  
Qy 2401 TTATTTGAGCCAGTTTCTTCAGAAATTTGTAGAAAGAACCCATAAGAAATTTAGTGTATTC 2460  
Db 2401 TTATTTGAGCCAGTTTCTTCAGAAATTTGTAGAAAGAACCCATAAGAAATTTAGTGTATTC 2460  
Qy 2461 AGAAGACATACAGAAAGAGGTGAAATGTTGAGGCTTTTGAGAAATTTGTCAITTTGACA 2520  
Db 2461 AGAAGACATACAGAAAGAGGTGAAATGTTGAGGCTTTTGAGAAATTTGTCAITTTGACA 2520  
Qy 2521 TATATGAGGTCAGATTTACTGCTTACTTGCACAGTGAAGAAAGAGTACATTTGA 2580  
Db 2521 TATATGAGGTCAGATTTACTGCTTACTTGCACAGTGAAGAAAGAGTACATTTGA 2580  
Qy 2581 TGAATATTTCTTTGTGGACTCTGCCACCTTCTGATGGTTTGCATCTATATATGACACA 2640  
Db 2581 TGAATATTTCTTTGTGGACTCTGCCACCTTCTGATGGTTTGCATCTATATATGACACA 2640  
Qy 2641 GAGTCTCAGAAATAGATGAAATGTTTGAAGCAGAAAGAAATGATTTGGCATTGTCACAGT 2700  
Db 2641 GAGTCTCAGAAATAGATGAAATGTTTGAAGCAGAAAGAAATGATTTGGCATTGTCACAGT 2700  
Qy 2701 TAGATATACACTTTTGTATGTTTGTACAGTAGAAGAAATTTATCAATTTTGGCTTCAATCA 2760  
Db 2701 TAGATATACACTTTTGTATGTTTGTACAGTAGAAGAAATTTATCAATTTTGGCTTCAATCA 2760  
Qy 2761 AAGGGATACCAGCCAAATATATACAGAAAGTGCAGAAAGGTTTACTAGATTTAGACA 2820  
Db 2761 AAGGGATACCAGCCAAATATATACAGAAAGTGCAGAAAGGTTTACTAGATTTAGACA 2820  
Qy 2821 TGCAGACTATCAAGATAACCAAGCTAAAAATTTAAGTGTGGTCAAAAAAGAAAGCTGT 2880  
Db 2821 TGCAGACTATCAAGATAACCAAGCTAAAAATTTAAGTGTGGTCAAAAAAGAAAGCTGT 2880  
Qy 2881 CATTAGGAATTTGCTTTTGGGAACCAAGACTACTGCTGTAGATGAACCAACAGCTG 2940  
Db 2881 CATTAGGAATTTGCTTTTGGGAACCAAGACTACTGCTGTAGATGAACCAACAGCTG 2940  
Qy 2941 GAATGGACCCCTGTTCTCGCATATTTGTATGGAATCTTTTAAATAACAGAAAGCCAAATC 3000  
Db 2941 GAATGGACCCCTGTTCTCGCATATTTGTATGGAATCTTTTAAATAACAGAAAGCCAAATC 3000  
Qy 3001 GGGTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATCTTTGACATAGGAAAG 3060  
Db 3001 GGGTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATCTTTGACATAGGAAAG 3060

Qy 3061 CTGTGATATCACAAGGAATGCTGAATGTGTGGTTCTTCAATGTTCCTCAAAAGTAAAT 3120  
Db 3061 CTGTGATATCACAAGGAATGCTGAATGTGTGGTTCTTCAATGTTCCTCAAAAGTAAAT 3120  
Qy 3121 GGGGGATCGGCTACCCGCTGAGCATGTACATAGACAAATATTGTGCCACAGAAATCTCTTT 3180  
Db 3121 GGGGGATCGGCTACCCGCTGAGCATGTACATAGACAAATATTGTGCCACAGAAATCTCTTT 3180  
Qy 3181 CTTTCACTGGTTAAACAACATATACCTGGAGCTACTTTTATACACAGAAATGACCAACAC 3240  
Db 3181 CTTTCACTGGTTAAACAACATATACCTGGAGCTACTTTTATACACAGAAATGACCAACAC 3240  
Qy 3241 TTGTGTATAGCTTGCCTTTTCAAGGACATGGAACAAATTTTTCAGGTTTGTCTTCTGCCCTAG 3300  
Db 3241 TTGTGTATAGCTTGCCTTTTCAAGGACATGGAACAAATTTTTCAGGTTTGTCTTCTGCCCTAG 3300  
Qy 3301 ACAGTCATTTCAAAATTTGGGTGTCATTTCTTATGTGTGTTCCATGACGACTTTTGGAAAGACG 3360  
Db 3301 ACAGTCATTTCAAAATTTGGGTGTCATTTCTTATGTGTGTTCCATGACGACTTTTGGAAAGACG 3360  
Qy 3361 TATTTTAAAGCTTAGAAGTTTGAAGCAGAAATTTGACCAAGCAGATTTATGTGTATTTACTC 3420  
Db 3361 TATTTTAAAGCTTAGAAGTTTGAAGCAGAAATTTGACCAAGCAGATTTATGTGTATTTACTC 3420  
Qy 3421 AGCAGCCACTTGGAGGAAGAAATGGATTTCAAAATCTTTTGTATGAAATGGAAACAGAGCTTAC 3480  
Db 3421 AGCAGCCACTTGGAGGAAGAAATGGATTTCAAAATCTTTTGTATGAAATGGAAACAGAGCTTAC 3480  
Qy 3481 TTATTTCTTTGAAACCAAGGCTTCTAGTGAAGCAACATGAGCCTTTGGAAACACACAGA 3540  
Db 3481 TTATTTCTTTGAAACCAAGGCTTCTAGTGAAGCAACATGAGCCTTTGGAAACACACAGA 3540  
Qy 3541 TGTATACAAATACCAAGTTTCAATTTCTTACCTTTGAAACCGTGAAGTAAATCAGTGAGAT 3600  
Db 3541 TGTATACAAATACCAAGTTTCAATTTCTTACCTTTGAAACCGTGAAGTAAATCAGTGAGAT 3600  
Qy 3601 CAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTGTTCCTGTTTCATC 3660  
Db 3601 CAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTGTTCCTGTTTCATC 3660  
Qy 3661 ACTTTTTAAAAATGCTGTGGTTCCCATCAAACTTGTTCAGACTTTATATTTTCTAAAAAC 3720  
Db 3661 ACTTTTTAAAAATGCTGTGGTTCCCATCAAACTTGTTCAGACTTTATATTTTCTAAAAAC 3720  
Qy 3721 CTGGACACAAACACATAAATACAAACCAAGTCTGCTTCTTCAAAATTTCTGCTGACTCAG 3780  
Db 3721 CTGGACACAAACACATAAATACAAACCAAGTCTGCTTCTTCAAAATTTCTGCTGACTCAG 3780  
Qy 3781 ATATCAGTCAATCTTATTTAGCTTTTTCACAAAGCCAGAAACATAATGTTGACGATGATTAATG 3840  
Db 3781 ATATCAGTCAATCTTATTTAGCTTTTTCACAAAGCCAGAAACATAATGTTGACGATGATTAATG 3840  
Qy 3841 ACAGTCACTATGATCCGCTGCTCCCATAGTGGCGCTTTTAAATGTTGATGATGATTCAGAAA 3900  
Db 3841 ACAGTCACTATGATCCGCTGCTCCCATAGTGGCGCTTTTAAATGTTGATGATGATTCAGAAA 3900  
Qy 3901 AGGACTATGTTTTTGGCAGCTGTTTTCAACAGTACTATGTTTTTCTTACCTATATTTAG 3960  
Db 3901 AGGACTATGTTTTTGGCAGCTGTTTTCAACAGTACTATGTTTTTCTTACCTATATTTAG 3960  
Qy 3961 TGAATATCAATTTAGTAACTACTTCTTTATCAATTTAAATGTGACTGAAACCATCCAGATCT 4020  
Db 3961 TGAATATCAATTTAGTAACTACTTCTTTATCAATTTAAATGTGACTGAAACCATCCAGATCT 4020  
Qy 4021 GGAGTACCCCAATCTTTTCAAGAAATTTACTGATATAGTTTTTAAAAATGAGCTGATTTTTC 4080  
Db 4021 GGAGTACCCCAATCTTTTCAAGAAATTTACTGATATAGTTTTTAAAAATGAGCTGATTTTTC 4080  
Qy 4081 AAGCAGCTTTGCTTGAATCAATTTGTTACTGCAATGCACTTACTTTGCCATGGAAGATG 4140  
Db 4081 AAGCAGCTTTGCTTGAATCAATTTGTTACTGCAATGCACTTACTTTGCCATGGAAGATG 4140

Qy	4141	CAGAGATCATAGATCAAAAGCTTATACCTAACTTTTCCAGTCTTTTGGCCATCTG	4200
Db	4141	CAGAGATCATAGATCAAAAGCTTATACCTAACTTTTGGCCATCTG	4200
Qy	4201	CATATTGGATTGGACAAAGCTGTTGGATATACCCCTTATTTTATCATCTTATTTGA	4260
Db	4201	CATATTGGATTGGACAAAGCTGTTGGATATACCCCTTATTTTATCATCTTATTTGA	4260
Qy	4261	TGCTAGGAAGCTTACTGGCAATTTTATGGAATATATTTTATCTGTAAGTTCCTTG	4320
Db	4261	TGCTAGGAAGCTTACTGGCAATTTTATGGAATATATTTTATCTGTAAGTTCCTTG	4320
Qy	4321	CTGTGGTTTTTGGCTTATTTGGTTATGTTCCATCAGTTATTTCTGTTCACTTATTTGCTT	4380
Db	4321	CTGTGGTTTTTGGCTTATTTGGTTATGTTCCATCAGTTATTTCTGTTCACTTATTTGCTT	4380
Qy	4381	CTTTCACCTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATCTATCTGTGG	4440
Db	4381	CTTTCACCTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATCTATCTGTGG	4440
Qy	4441	CAGGTTGNCCTTATTTGCAATCACTGAATTAATCTTTTATGGGATACACAATTTGCAA	4500
Db	4441	CAGGTTGNCCTTATTTGCAATCACTGAATTAATCTTTTATGGGATACACAATTTGCAA	4500
Qy	4501	CTATCTTCAATATGCTTTTGTATCATCATTTCCAATCTATCCACTTTCTAGGTTGCTGA	4560
Db	4501	CTATCTTCAATATGCTTTTGTATCATCATTTCCAATCTATCCACTTTCTAGGTTGCTGA	4560
Qy	4561	TTTCTTTTCAAAAGATTTCTTGGAAAGATGTAAGAAATTTGGAACCACTTATTAATCCAT	4620
Db	4561	TTTCTTTTCAAAAGATTTCTTGGAAAGATGTAAGAAATTTGGAACCACTTATTAATCCAT	4620
Qy	4621	GGGATAGGCTTTCAGTAGCTGTTATATCGCTTACCTGCAAGTGTACTGTGATTTTCC	4680
Db	4621	GGGATAGGCTTTCAGTAGCTGTTATATCGCTTACCTGCAAGTGTACTGTGATTTTCC	4680
Qy	4681	TCCTTACAATACTATGAAAGAAATATGGAGGAGATCAATAAGAAAGATCCCTTTTCA	4740
Db	4681	TCCTTACAATACTATGAAAGAAATATGGAGGAGATCAATAAGAAAGATCCCTTTTCA	4740
Qy	4741	GAAACCTTTTCAAGAGCTTAAAGATAGGAAGCTTCCAGAACCAACAGACAATGAGGATG	4800
Db	4741	GAAACCTTTTCAAGAGCTTAAAGATAGGAAGCTTCCAGAACCAACAGACAATGAGGATG	4800
Qy	4801	AAGATGAAGTGTCAAGCTTAAAGACTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTT	4860
Db	4801	AAGATGAAGTGTCAAGCTTAAAGACTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTT	4860
Qy	4861	GTGAGGAAACCATCCATTTATGCTCAGCAATTTGCAATGAAGATATGATGACAAGAAAG	4920
Db	4861	GTGAGGAAACCATCCATTTATGCTCAGCAATTTGCAATGAAGATATGATGACAAGAAAG	4920
Qy	4921	ATTTTCTTTTCAAGAAAGTAAAGAAAGTGGCAACTAAATACATCTCTTCTGTGTA	4980
Db	4921	ATTTTCTTTTCAAGAAAGTAAAGAAAGTGGCAACTAAATACATCTCTTCTGTGTA	4980
Qy	4981	AAAAAGAGAGATCTTAGGACTATTTGGTCCAAATGGTGTGCAAGCAACAAATTTA	5040
Db	4981	AAAAAGAGAGATCTTAGGACTATTTGGTCCAAATGGTGTGCAAGCAACAAATTTA	5040
Qy	5041	ATATTCTGGTTGGTATATTGAACCACTTTCAGCCAGGATTTTATAGGAGATTTCTT	5100
Db	5041	ATATTCTGGTTGGTATATTGAACCACTTTCAGCCAGGATTTTATAGGAGATTTCTT	5100
Qy	5101	CAGACAACTGAAGATGATTTCACTGAAGTGTATGGGTTACTGTCTCAGATAAAC	5160
Db	5101	CAGACAACTGAAGATGATTTCACTGAAGTGTATGGGTTACTGTCTCAGATAAAC	5160
Qy	5161	CTTTGTGGCAGATCTACATTTCCAGGAACATTTTGAATTTTATGGAGCTGTCAAAGGA	5220
Db	5161	CTTTGTGGCAGATCTACATTTCCAGGAACATTTTGAATTTTATGGAGCTGTCAAAGGA	5220
Qy	5221	TGAGTGAAGTGAATGAAAGAGTCTAATAGTGAATTAACATGCACTTTGATTTAAAG	5280

Db	5221	TGAGTGAAGTGAATGAAAGAGTCTAATAGTGAATTAACATGCACTTTGATTTAAAG	5280
Qy	5281	AACATCTTTCAGAGACTGTAAGAAACTTACCTGAGGAATCAAAAGAAAGTTGTGTTTG	5340
Db	5281	AACATCTTTCAGAGACTGTAAGAAACTTACCTGAGGAATCAAAAGAAAGTTGTGTTTG	5340
Qy	5341	CTTAAAGTATGCTAGGGAATCTCTCAGATTTACTTTGCTAGATGAACCATCTACAGGTATG	5400
Db	5341	CTTAAAGTATGCTAGGGAATCTCTCAGATTTACTTTGCTAGATGAACCATCTACAGGTATG	5400
Qy	5401	ATCCAAAGCCAAACAGACATGTTGGCGAGCAATTTGCAACTGCAATTTAAACAGAAAGC	5460
Db	5401	ATCCAAAGCCAAACAGACATGTTGGCGAGCAATTTGCAACTGCAATTTAAACAGAAAGC	5460
Qy	5461	GGGCTGCTTCTGACCACTCACTATATGAGGAGGAGAGGCTCTCTGATCGAGTAG	5520
Db	5461	GGGCTGCTTCTGACCACTCACTATATGAGGAGGAGAGGCTCTCTGATCGAGTAG	5520
Qy	5521	CTATCATGCTCTCTGGGAGTTAAGATGTATCGGAAACAGTACAACTCTAAAGAGTAAT	5580
Db	5521	CTATCATGCTCTCTGGGAGTTAAGATGTATCGGAAACAGTACAACTCTAAAGAGTAAT	5580
Qy	5581	TTGGAAGAGGCTACTTTTGGAAATTTAAATTTGAAAGCTGGATAGAAAACCTAGAGTAG	5640
Db	5581	TTGGAAGAGGCTACTTTTGGAAATTTAAATTTGAAAGCTGGATAGAAAACCTAGAGTAG	5640
Qy	5641	ACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCGTCAGGAAAGTTTTT	5700
Db	5641	ACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCGTCAGGAAAGTTTTT	5700
Qy	5701	CTTCTATTTTGGCTTATATAAATTTCTAAGGAAGATGTTTCACTCCCTTTTCAAACTTTT	5760
Db	5701	CTTCTATTTTGGCTTATATAAATTTCTAAGGAAGATGTTTCACTCCCTTTTCAAACTTTT	5760
Qy	5761	TTAAGCTGGAAGAGCTAAAACATGCTTTTGGCAATTTGAAGATAATAGCTTTTCTCAAGCA	5820
Db	5761	TTAAGCTGGAAGAGCTAAAACATGCTTTTGGCAATTTGAAGATAATAGCTTTTCTCAAGCA	5820
Qy	5821	CATTGGAACAGGTTTCTAGAACTCACTAAGACAGAGGAGGAAGATAATAGTTCTG	5880
Db	5821	CATTGGAACAGGTTTCTAGAACTCACTAAGAAACAGAGAGGAGGAAGATAATAGTTCTG	5880
Qy	5881	GAACTTTAAACACACACTTTTGGTGGGAAACGAAACACAGAAAGATAGATTTTGA	5940
Db	5881	GAACTTTAAACACACACTTTTGGTGGGAAACGAAACACAGAAAGATAGATTTTGA	5940
Qy	5941	TTTGTATTGTTTGGTCTGCTTACTGGGACTTCTTTTCACTTAATTTTAACTTTG	6000
Db	5941	TTTGTATTGTTTGGTCTGCTTACTGGGACTTCTTTTCACTTAATTTTAACTTTG	6000
Qy	6001	TTTAAAGATTTTATTTGGAATGTTTCTGAGAACCAAGACGCACTTGAATTTTTC	6060
Db	6001	TTTAAAGATTTTATTTGGAATGTTTCTGAGAACCAAGACGCACTTGAATTTTTC	6060
Qy	6061	TAAAGCTCTTAAATGAAATGCTGTTGTTGTTTCTTTTCTTTTAAATAAAGCTAT	6120
Db	6061	TAAAGCTCTTAAATGAAATGCTGTTGTTGTTTCTTTTCTTTTAAATAAAGCTAT	6120
Qy	6121	GTATAAATTAAGTGAAGCTGCTATTTGTTTGAAGTATATTGAACATATATAGTTGTATG	6180
Db	6121	GTATAAATTAAGTGAAGCTGCTATTTGTTTGAAGTATATTGAACATATATAGTTGTATG	6180
Qy	6181	TCATCTTTTCCACCACTTCAAGAACAGTCTTCTGAAATTTGATTTAAAGGAAATTTAAT	6240
Db	6181	TCATCTTTTCCACCACTTCAAGAACAGTCTTCTGAAATTTGATTTAAAGGAAATTTAAT	6240
Qy	6241	AGAAATGTTTTTATTTTAAAGTTATCTTTAAGTTTATGCTCTTTTAAATAAGTACGTA	6300
Db	6241	AGAAATGTTTTTATTTTAAAGTTATCTTTAAGTTTATGCTCTTTTAAATAAGTACGTA	6300
Qy	6301	ATGTTTCCAACTTAAATAAAGAACTTAAATCATTAATGCTATAGAAAGATACATAAGC	6360
Db	6301	ATGTTTCCAACTTAAATAAAGAACTTAAATCATTAATGCTATAGAAAGATACATAAGC	6360



Db 781 ATTCTGGTCTACCTATTTTCCATGCTGCTTTTACTTTTTCAGAGCTGACAGATTGCTGCT 840  
Qy 841 CCATGCAATTCCTGCTCCAGTTTCCCTAAGAGAGACAGCTTGGAGTATGCTTAATCCATCTTAC 900  
Db 841 CCATGCAATTCCTGCTCCAGTTTCCCTAAGAGAGACAGCTTGGAGTATGCTTAATCCATCTTAC 900  
Qy 901 CTGGGACTGAAACAGCTGCTTATTTTGGCCGTTAAATAATTACATGCAAGTTTACTGCGTGGC 960  
Db 901 CTGGGACTGAAACAGCTGCTTATTTTGGCCGTTAAATAATTACATGCAAGTTTACTGCGTGGC 960  
Qy 961 TCCGGGTTGCTTGTGTTGTTTTCCTCTTTTAAATAGGTTTATTCAGAAAAACATGCTCCACTG 1020  
Db 961 TCCGGGTTGCTTGTGTTGTTTTCCTCTTTTAAATAGGTTTATTCAGAAAAACATGCTCCACTG 1020  
Qy 1021 CAATTAGGAGGTAGGAGTTTGGAGACAGACAGACACTTCTACTGGAAGAAATTAATTAA 1080  
Db 1021 CAATTAGGAGGTAGGAGTTTGGAGACAGACAGACACTTCTACTGGAAGAAATTAATTAA 1080  
Qy 1081 TTAATGTCAGAAACCAAAAGAGTAGTGTTCAGGAAATTCCTTTTCCACTATTTTATTTTAT 1140  
Db 1081 TTAATGTCAGAAACCAAAAGAGTAGTGTTCAGGAAATTCCTTTTCCACTATTTTATTTTAT 1140  
Qy 1141 TTTGGTTAATTAATTAATAGCATGATGCATCCAAATAAGAAATATGAAGAGTGCCTTAATA 1200  
Db 1141 TTTGGTTAATTAATTAATAGCATGATGCATCCAAATAAGAAATATGAAGAGTGCCTTAATA 1200  
Qy 1201 TAGAATCAATCCTATGGAACAAGTTTACTCTCTTCTTAATCTAAATCTTGAATATACTCCAG 1260  
Db 1201 TAGAATCAATCCTATGGAACAAGTTTACTCTCTTCTTAATCTAAATCTTGAATATACTCCAG 1260  
Qy 1261 TGACTAATATTAACAAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTACCTGATGTCA 1320  
Db 1261 TGACTAATATTAACAAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTACCTGATGTCA 1320  
Qy 1321 TAAATCTGAAGAAATATACAAATGAAAGAAATGTTTAAATCCAGTCTCTCTTAAGCCGA 1380  
Db 1321 TAAATCTGAAGAAATATACAAATGAAAGAAATGTTTAAATCCAGTCTCTCTTAAGCCGA 1380  
Qy 1381 GCAACTTTGAGTGCTGTTTCAAGACCTCCATGCTTCTATGAACTTCCGTTTTCCTG 1440  
Db 1381 GCAACTTTGAGTGCTGTTTCAAGACCTCCATGCTTCTATGAACTTCCGTTTTCCTG 1440  
Qy 1441 ATATGATTCAGATATCTCTATTTATATGATTAAGAGCTGCTGTTTCAAAATCATGTG 1500  
Db 1441 ATATGATTCAGATATCTCTATTTATATGATTAAGAGCTGCTGTTTCAAAATCATGTG 1500  
Qy 1501 AGGCTGCTCAGTACTGCTCCTCAGGTTTTCAGTTTTTACAAGCATCCATAGATGCTGCCA 1560  
Db 1501 AGGCTGCTCAGTACTGCTCCTCAGGTTTTCAGTTTTTACAAGCATCCATAGATGCTGCCA 1560  
Qy 1561 TTATACAGTTGAAGACCAATGTTTCTCTTTGGAGAGCTGGAGTCAACTAAAGCTGTGA 1620  
Db 1561 TTATACAGTTGAAGACCAATGTTTCTCTTTGGAGAGCTGGAGTCAACTAAAGCTGTGA 1620  
Qy 1621 TTATGGAGAACTGCTGTTGTAGAAATAGATACCTTTTCCGAGGAGTAAATTTAATAT 1680  
Db 1621 TTATGGAGAACTGCTGTTGTAGAAATAGATACCTTTTCCGAGGAGTAAATTTAATAT 1680  
Qy 1681 ACCTAGTTATAGCAATTTTCCACTTTTGGATACCTTTTGGCAATTCATATCTGACGAGAA 1740  
Db 1681 ACCTAGTTATAGCAATTTTCCACTTTTGGATACCTTTTGGCAATTCATATCTGACGAGAA 1740  
Qy 1741 AAGAAAAAATAAAGAAATTTTAAAGATAATAGGACTTCATGATACGTGCTTTTGGC 1800  
Db 1741 AAGAAAAAATAAAGAAATTTTAAAGATAATAGGACTTCATGATACGTGCTTTTGGC 1800  
Qy 1801 TTTCTGGTTCCTCTATATACAGTTTAAATTTTCTTATGCTCCCTCTTTATGGCAGTCA 1860  
Db 1801 TTTCTGGTTCCTCTATATACAGTTTAAATTTTCTTATGCTCCCTCTTTATGGCAGTCA 1860  
Qy 1861 TTGCGACAGCTTCTTTGTTATTTCTTCAAAAGTAGCAGCATGTTGATATTTCTGCTTTTTT 1920  
Db 1861 TTGCGACAGCTTCTTTGTTATTTCTTCAAAAGTAGCAGCATGTTGATATTTCTGCTTTTTT 1920

Qy 1921 TCCTTTATGGATTATCATCTGTATTTTTTGTCTTTAATGCTGACACCTCTTTTTAAAAAAT 1980  
Db 1921 TCCTTTATGGATTATCATCTGTATTTTTTGTCTTTAATGCTGACACCTCTTTTTAAAAAAT 1980  
Qy 1981 CAAAACATGTGGGAATAGTTCGAAATTTTTTGTACTGTGGCTTTTGGATTTATGTGCCCCTTA 2040  
Db 1981 CAAAACATGTGGGAATAGTTCGAAATTTTTTGTACTGTGGCTTTTGGATTTATGTGCCCCTTA 2040  
Qy 2041 TGATTAATCCTCATAGAAAGTTTTCCTCCAAATCGTTAGTGTGGCTTTTTCAGTCTCTTCTGTC 2100  
Db 2041 TGATTAATCCTCATAGAAAGTTTTCCTCCAAATCGTTAGTGTGGCTTTTTCAGTCTCTTCTGTC 2100  
Qy 2101 ACTGTACTTTTGTGATTTGCTATTTGCACAGGTTCATGCAATTTAGAGATTTAATGAAGGTG 2160  
Db 2101 ACTGTACTTTTGTGATTTGCTATTTGCACAGGTTCATGCAATTTAGAGATTTAATGAAGGTG 2160  
Qy 2161 CTTTCAATTTTCAAAATTTGACTGCGAGGCCCATATCTCTTAATTTATTAACAATTTATCATGCTCA 2220  
Db 2161 CTTTCAATTTTCAAAATTTGACTGCGAGGCCCATATCTCTTAATTTATTAACAATTTATCATGCTCA 2220  
Qy 2221 CACTTAATAGTATATTTCTATGTCCTCTTGGCTGTCTATCTTGAAGCCTTTCATATTTGCTCAGGG 2280  
Db 2221 CACTTAATAGTATATTTCTATGTCCTCTTGGCTGTCTATCTTGAAGCCTTTCATATTTGCTCAGGG 2280  
Qy 2281 AATTTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCTTTCATATTTGCTCAGAGAGTA 2340  
Db 2281 AATTTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCTTTCATATTTGCTCAGAGAGTA 2340  
Qy 2341 AAAGAAATTTATGAGGAGTTATCAGAGGCCAATGTTTAAATGGAATAATTTAGTTTTAGTGAAA 2400  
Db 2341 AAAGAAATTTATGAGGAGTTATCAGAGGCCAATGTTTAAATGGAATAATTTAGTTTTAGTGAAA 2400  
Qy 2401 TTATTTGAGCCAGTTTCTTCAGAAATTTGTAGAAAGAGCCATAGAAATTTAGTGTATTC 2460  
Db 2401 TTATTTGAGCCAGTTTCTTCAGAAATTTGTAGAAAGAGCCATAGAAATTTAGTGTATTC 2460  
Qy 2461 AGAAGACATACAGAAAGAGGTGAAATGTGGAGCTTTTGAGAAATTTTGTCAATTTTGACA 2520  
Db 2461 AGAAGACATACAGAAAGAGGTGAAATGTGGAGCTTTTGAGAAATTTTGTCAATTTTGACA 2520  
Qy 2521 TATATGAGGGTCAGATTTACTGCTTTTCTGGCCACAGTGGAAACAGGAAAGAGTACATTCGA 2580  
Db 2521 TATATGAGGGTCAGATTTACTGCTTTTCTGGCCACAGTGGAAACAGGAAAGAGTACATTCGA 2580  
Qy 2581 TGAATATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATATATGACACA 2640  
Db 2581 TGAATATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATATATGACACA 2640  
Qy 2641 GAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAAAATGATTTGGCATTTTGTCCACAGT 2700  
Db 2641 GAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAAAATGATTTGGCATTTTGTCCACAGT 2700  
Qy 2701 TAGATATACACTTTGATGTTTGTACAGTAGAAGAAAAATTTATCAATTTTGGCTTCAATCA 2760  
Db 2701 TAGATATACACTTTGATGTTTGTACAGTAGAAGAAAAATTTATCAATTTTGGCTTCAATCA 2760  
Qy 2761 AAGGATACCGCCCAACAAATATTAACAAGTGCAGAAAGTTTTTACTAGATTTTAGACA 2820  
Db 2761 AAGGATACCGCCCAACAAATATTAACAAGTGCAGAAAGTTTTTACTAGATTTTAGACA 2820  
Qy 2821 TGCAGACTATCAAGATTAACCAAGCTTAAAGTTTAAAGTTTGGTCAAAAAAGAAAGCTGT 2880  
Db 2821 TGCAGACTATCAAGATTAACCAAGCTTAAAGTTTAAAGTTTGGTCAAAAAAGAAAGCTGT 2880  
Qy 2881 CATTAGGAATTCGCTGTTCTTTGGGAAACCCAAAGATGCTGCTGATGAAACCAAGCTGT 2940  
Db 2881 CATTAGGAATTCGCTGTTCTTTGGGAAACCCAAAGATGCTGCTGATGAAACCAAGCTGT 2940  
Qy 2941 GAATGGAACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAATACAGAAAGCCAAATC 3000  
Db 2941 GAATGGAACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAATACAGAAAGCCAAATC 3000

QY 3001 GGGTCAGAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATCTTTCAGATAGGAAAG 3060  
DB 3001 GGGTCAGAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATCTTTCAGATAGGAAAG 3060  
QY 3061 CTGTGATATCAACAGGAATGCTGAAATGCTGTGGTCTTCAATGTTCTTCAAAAGTAAAT 3120  
DB 3061 CTGTGATATCAACAGGAATGCTGAAATGCTGTGGTCTTCAATGTTCTTCAAAAGTAAAT 3120  
QY 3121 GGGGGATCGGCTACCGCTGAGCTGTATACATAGACAAATATGTGGCCACAGAAATCTCTTT 3180  
DB 3121 GGGGGATCGGCTACCGCTGAGCTGTATACATAGACAAATATGTGGCCACAGAAATCTCTTT 3180  
QY 3181 CTTTCACTGTTTAAACACATATACCTGGAGCTACTTTATTAACAACAGATGACCAAC 3240  
DB 3181 CTTTCACTGTTTAAACACATATACCTGGAGCTACTTTATTAACAACAGATGACCAAC 3240  
QY 3241 TTGTGTATAGCTTGCCTTTCAAGACATGGAACAAATTTTCAGGTTTCTGCCCCTAG 3300  
DB 3241 TTGTGTATAGCTTGCCTTTCAAGACATGGAACAAATTTTCAGGTTTCTGCCCCTAG 3300  
QY 3301 ACAGTCATTTCAAAATTTGGGTGTCATTTCTTATGGTGTTCATGACGACTTTTGGAGACG 3360  
DB 3301 ACAGTCATTTCAAAATTTGGGTGTCATTTCTTATGGTGTTCATGACGACTTTTGGAGACG 3360  
QY 3361 TATTTTTAAAGCTAGAGTTGAACGAGAAATGACCAAGCAGATTAAGTGTATTTACTC 3420  
DB 3361 TATTTTTAAAGCTAGAGTTGAACGAGAAATGACCAAGCAGATTAAGTGTATTTACTC 3420  
QY 3421 AGCAGCCACTGGAGGAGAAATGATTCAAATCTTTTGTATGAATGGAACAGAGCTTAC 3480  
DB 3421 AGCAGCCACTGGAGGAGAAATGATTCAAATCTTTTGTATGAATGGAACAGAGCTTAC 3480  
QY 3481 TTAATTTCTTGAACCAAGGCTTCTCTAGTGAACCATGAGCCTTTTGGAAACAAACAGA 3540  
DB 3481 TTAATTTCTTGAACCAAGGCTTCTCTAGTGAACCATGAGCCTTTTGGAAACAAACAGA 3540  
QY 3541 TGATACAAATAGCAAGTTTCATTTCTTACCTTGAACCGTGAAGTAAATTCAGTGAGAT 3600  
DB 3541 TGATACAAATAGCAAGTTTCATTTCTTACCTTGAACCGTGAAGTAAATTCAGTGAGAT 3600  
QY 3601 CAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTTTTTGGTTCATC 3660  
DB 3601 CAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTTTTTGGTTCATC 3660  
QY 3661 ACTCTTTTAAATATGCTGTGTTTCCCATCAAACTGTTTCAGACTTATATTTTCTAAAC 3720  
DB 3661 ACTCTTTTAAATATGCTGTGTTTCCCATCAAACTGTTTCAGACTTATATTTTCTAAAC 3720  
QY 3721 CTGGAGACAAACCAATAAACAAGTCTGCTTCTTCAAAATTCGTGACTCAG 3780  
DB 3721 CTGGAGACAAACCAATAAACAAGTCTGCTTCTTCAAAATTCGTGACTCAG 3780  
QY 3781 ATATCAGTGATCTTATAGCTTTTTCACAGCCAGAACATAATGGTGACGATTAATG 3840  
DB 3781 ATATCAGTGATCTTATAGCTTTTTCACAGCCAGAACATAATGGTGACGATTAATG 3840  
QY 3841 ACAGTGACTATGATCCGTGGCTCCCATAGTGGGCTTTAAATGTGATGCAITTCAGAAA 3900  
DB 3841 ACAGTGACTATGATCCGTGGCTCCCATAGTGGGCTTTAAATGTGATGCAITTCAGAAA 3900  
QY 3901 AGGACTATGTTTTGAGCTGTTTTTCAAGTACTATGTTTATCTTTACCTATATTAG 3960  
DB 3901 AGGACTATGTTTTGAGCTGTTTTTCAAGTACTATGTTTATCTTTACCTATATTAG 3960  
QY 3961 TGAATATCATTAGTAACTACTATCTTTTATCAATGATGCTGAAACCAATCCAGATCT 4020  
DB 3961 TGAATATCATTAGTAACTACTATCTTTTATCAATGATGCTGAAACCAATCCAGATCT 4020  
QY 4021 GGAGTACCCCATTTTCAAGAAATTAATGATAGTTTTTAAATTTGAGCTGTATTTTC 4080  
DB 4021 GGAGTACCCCATTTTCAAGAAATTAATGATAGTTTTTAAATTTGAGCTGTATTTTC 4080  
QY 4081 AAGCAGCTTGTGGAATCATTTGTTACTGCAATGCCACTTACTTTTGGCAATGGAATG 4140

DB 4081 AAGCAGCTTGTGGAATCATTTGTTACTGCAATGCCACTTACTTTGCCATGGAATG 4140  
QY 4141 CAGAGATCATAGATCAAAAGCTTATACCTCAACTTAACTTTTTCAGGCTTTTGGCCATCTG 4200  
DB 4141 CAGAGATCATAGATCAAAAGCTTATACCTCAACTTAACTTTTTCAGGCTTTTGGCCATCTG 4200  
QY 4201 CATATTGGATTGGACAGCTGTTGTGATATCCCTTATTTTATCATTTTATTTTGA 4260  
DB 4201 CATATTGGATTGGACAGCTGTTGTGATATCCCTTATTTTATCATTTTATTTTGA 4260  
QY 4261 TGCTAGGAAGCTTACTGGCAATTTTCAATGGAATATATTTTATCTGTAAGTTCCTTG 4320  
DB 4261 TGCTAGGAAGCTTACTGGCAATTTTCAATGGAATATATTTTATCTGTAAGTTCCTTG 4320  
QY 4321 CTGTGGTTTTTTCGCTTATTTGGTATTTGTTCCATCAGTTATCTGTTCACATTAATGCTT 4380  
DB 4321 CTGTGGTTTTTTCGCTTATTTGGTATTTGTTCCATCAGTTATTTCTGTTCACATTAATGCTT 4380  
QY 4381 CTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTTCATTTATCTATTTCTGTG 4440  
DB 4381 CTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTTCATTTATCTATTTCTGTG 4440  
QY 4441 CAGCGTTGNCCTTGTATTTGCAATCACTGAAATACTTTCTTTTATGGGATACCAATTTGCAA 4500  
DB 4441 CAGCGTTGNCCTTGTATTTGCAATCACTGAAATACTTTCTTTTATGGGATACCAATTTGCAA 4500  
QY 4501 CTATTTCTTCAATTTATGCTTTTGTATCATCATTTCCAATCTATCCATTTCTAGGTTGCTGA 4560  
DB 4501 CTATTTCTTCAATTTATGCTTTTGTATCATCATTTCCAATCTATCCATTTCTAGGTTGCTGA 4560  
QY 4561 TTTCTTTTCAATAAGATTTCTTGGAGAAATGTAGAAAATTTGGACACCTTATAATCCAT 4620  
DB 4561 TTTCTTTTCAATAAGATTTCTTGGAGAAATGTAGAAAATTTGGACACCTTATAATCCAT 4620  
QY 4621 GGGATAGGCTTTTCAAGTGTGTTTATATCGCTTTACCTGCAAGTGTGCTGATTTTCC 4680  
DB 4621 GGGATAGGCTTTTCAAGTGTGTTTATATCGCTTTACCTGCAAGTGTGCTGATTTTCC 4680  
QY 4681 TCTTACATACTATGAGAAAATATTTGGAGGAGATCAATTAAGAAAGATCCCTTTTCA 4740  
DB 4681 TCTTACATACTATGAGAAAATATTTGGAGGAGATCAATTAAGAAAGATCCCTTTTCA 4740  
QY 4741 GAAACCTTTCAAGAGTCTTAAATATAGGAAGCTTCCAGAACCCAGACAAATGAGATG 4800  
DB 4741 GAAACCTTTCAAGAGTCTTAAATATAGGAAGCTTCCAGAACCCAGACAAATGAGATG 4800  
QY 4801 AAGATGAAGATGTCAAAGCTGAAAGCTAAAGGTCAAAGAGCTGATGGGTTGCCAGTGT 4860  
DB 4801 AAGATGAAGATGTCAAAGCTGAAAGCTAAAGGTCAAAGAGCTGATGGGTTGCCAGTGT 4860  
QY 4861 GTGAGGAGAAACCATCCATTTATGTTGTCAGCAATTTGCAATTAAGAAATATGATGACAGAAAG 4920  
DB 4861 GTGAGGAGAAACCATCCATTTATGTTGTCAGCAATTTGCAATTAAGAAATATGATGACAGAAAG 4920  
QY 4921 ATTTTCTTTTCAAGAAAGTAAAGAGTGGCAACTTAAATACATCTCTTCTGTGTCGA 4980  
DB 4921 ATTTTCTTTTCAAGAAAGTAAAGAGTGGCAACTTAAATACATCTCTTCTGTGTCGA 4980  
QY 4981 AAAAAAGGAGATCTTAGGACTTATTTGGGTCCAAATTTGGTCTGGCAAAAGCAATTTATTA 5040  
DB 4981 AAAAAAGGAGATCTTAGGACTTATTTGGGTCCAAATTTGGTCTGGCAAAAGCAATTTATTA 5040  
QY 5041 ATATTTCTGGTGGTATTTGAAACCACTTCHAGGCCAGGTATTTTTAGGAGATTAATCTTT 5100  
DB 5041 ATATTTCTGGTGGTATTTGAAACCACTTCHAGGCCAGGTATTTTTAGGAGATTAATCTTT 5100  
QY 5101 CAGAGACAAGTGAAGATGATGATTTCACTGAAGTGTATGGGTTACTGCTCAGATAAAC 5160  
DB 5101 CAGAGACAAGTGAAGATGATGATTTCACTGAAGTGTATGGGTTACTGCTCAGATAAAC 5160  
QY 5161 CTTTGTGCCAGATACTACATTTGAGGAAACATTTTGAATTTTATGGAGCTGTCAAGGAA 5220







Qy	2521	TATATAGAGGTCAGATTACTGCTTACTTGGCCACAGTCAGTGAACAGGAAAGAGTACATTGA	2580		Db	3828	CAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTTTGGTTTCATC	3887	
Db	2748	TATATAGAGGTCAGATTACTGCTTACTTGGCCACAGTCAGTGAACAGGAAAGAGTACATTGA	2807		Qy	3661	ACTCTTTTAAATAACTGCTGGTTCCTCAATCAAACTTTGTTCCAGACTTATATATTTCTAAAC	3720	
Qy	2581	TGAATATTTCTTTGGGACTCTGCCACTCTCTGATGGTTTGGATCTATATATATGAGACACA	2640		Db	3888	ACTCTTTTAAATAACTGCTGGTTCCTCAATCAAACTTTGTTCCAGACTTATATATTTCTAAAC	3947	
Db	2808	TGAATATTTCTTTGGGACTCTGCCACTCTCTGATGGTTTGGATCTATATATATGAGACACA	2867		Qy	3721	CTGGAGACAAACCAATATAAATAAATAAAGTCTGCTTCTTCAAAATTCGCTGACTCAG	3780	
Qy	2641	GAGTCTCAGAAATAGATGAATATTTGAAGCAAGAAATAATGATGGCTTTGTCACAGT	2700		Db	3948	CTGGAGACAAACCAATATAAATAAATAAAGTCTGCTTCTTCAAAATTCGCTGACTCAG	4007	
Db	2868	GAGTCTCAGAAATAGATGAATATTTGAAGCAAGAAATAATGATGGCTTTGTCACAGT	2927		Qy	3781	ATATCAGTGATCTTATAGCTTTTTCACAGCCAGAACATAATATGTTGAGCATGATTAATG	3840	
Qy	2701	TAGATATACACTTTGATGTTTTGACAGTAGAAGAAATTTATCAATTTTGGCTTTCAATCA	2760		Db	4008	ATATCAGTGATCTTATAGCTTTTTCACAGCCAGAACATAATATGTTGAGCATGATTAATG	4067	
Db	2928	TAGATATACACTTTGATGTTTTGACAGTAGAAGAAATTTATCAATTTTGGCTTTCAATCA	2987		Qy	3841	ACAGTGACTATGATATCCGTGGCTCCCATAGTGGCTTTAAATGTTGATGATGATCAGAAA	3900	
Qy	2761	AAGGGATACCAGGCCAACATATATAACAAGAGTGCAGAAAGTTTTTACTAGATTTTAGACA	2820		Db	4068	ACAGTGACTATGATATCCGTGGCTCCCATAGTGGCTTTAAATGTTGTTGATCAGAAA	4127	
Db	2988	AAGGGATACCAGGCCAACATATAATAACAAGAGTGCAGAAAGTTTTTACTAGATTTTAGACA	3047		Qy	3901	AGGACTATGTTTTTTCAGCTGTTTTCAACAGTACTATGTTTATTTCTTTTACCTATATTAG	3960	
Qy	2821	TGCAGACTATCAAGATAACCAAGCTTAAATAATTTAGTGGTCTCAAAAAGAAAGCTGT	2880		Db	4128	AGGACTATGTTTTTTCAGCTGTTTTCAACAGTACTATGTTTATTTCTTTTACCTATATTAG	4187	
Db	3048	TGCAGACTATCAAGATAACCAAGCTTAAATAATTTAGTGGTCTCAAAAAGAAAGCTGT	3107		Qy	3961	TGAATATCATTTAGTAACTACTATCTTTTATCATTTTAAATGTGACTGAAACCATCCAGATCT	4020	
Qy	2881	CATTAGGAATTTGCTGTTCTTTGGGAACCCAAAGATACCTGCTAGATGAACCAACAGCTG	2940		Db	4188	TGAATATCATTTAGTAACTACTATCTTTTATCATTTTAAATGTGACTGAAACCATCCAGATCT	4247	
Db	3108	CATTAGGAATTTGCTGTTCTTTGGGAACCCAAAGATACCTGCTAGATGAACCAACAGCTG	3167		Qy	4021	GGAGTACCCCATCTTTTCAAGAAATTAATGATATGTTTTTAAATTTGAGCTGATTTTC	4080	
Qy	2941	GAATGGACCCCTGTTCTCGACATATTGTTATGGAATCTTTTAAATAACAGAAAGCCAAATC	3000		Db	4248	GGAGTACCCCATCTTTTCAAGAAATTAATGATATGTTTTTAAATTTGAGCTGATTTTC	4307	
Db	3168	GAATGGACCCCTGTTCTCGACATATTGTTATGGAATCTTTTAAATAACAGAAAGCCAAATC	3227		Qy	4081	AAGCAGCTTTGCTTGGATCATTTGTTACTGCAATGCCACCTTACTTTTGGCATGGAAAATG	4140	
Qy	3001	GGGTGACAGTTTCAGTACTCATTTTCATGATGAAGCTGACATCTTTCAGATAGGAAAG	3060		Db	4308	AAGCAGCTTTGCTTGGATCATTTGTTACTGCAATGCCACCTTACTTTTGGCATGGAAAATG	4367	
Db	3228	GGGTGACAGTTTCAGTACTCATTTTCATGATGAAGCTGACATCTTTCAGATAGGAAAG	3287		Qy	4141	CAGAGAATCATTAAGATCAAAAGCTTATATCTAACTTTAAACTTTTTCAGGTCTTTTGGCATCTG	4200	
Qy	3061	CTGTGATATCAAGGAATGCTGAAATGTTGTTGTTCTTCAATGTTTCTTCAAAAGTAAAT	3120		Db	4368	CAGAGAATCATTAAGATCAAAAGCTTATATCTAACTTTAAACTTTTTCAGGTCTTTTGGCATCTG	4427	
Db	3288	CTGTGATATCAAGGAATGCTGAAATGTTGTTGTTCTTCAATGTTTCTTCAAAAGTAAAT	3347		Qy	4201	CATATTTGATTTGGACAAAGCTGTTGTTGATATCCCTTATTTTATCATTTCTTATTTTGA	4260	
Qy	3121	GGGGGATTCGGCTACCGCTGAGCATGATACAGAAATATTTGTCACAGAAATCTCTTT	3180		Db	4428	CATATTTGATTTGGACAAAGCTGTTGTTGATATCCCTTATTTTATCATTTCTTATTTTGA	4487	
Db	3348	GGGGGATTCGGCTACCGCTGAGCATGATACAGAAATATTTGTCACAGAAATCTCTTT	3407		Qy	4261	TGCTAGGAAGCTTACTGTCATTTCAATTTGGAATTAATTTTATATCTGTAAGTTTCTCTTG	4320	
Qy	3181	CTTCAGTGGTTTAAACACATATACCTGGAGCTACTTTTATACACAGAAATGACCAACAC	3240		Db	4488	TGCTAGGAAGCTTACTGTCATTTCAATTTGGAATTAATTTTATATCTGTAAGTTTCTCTTG	4547	
Db	3408	CTTCAGTGGTTTAAACACATATACCTGGAGCTACTTTTATACACAGAAATGACCAACAC	3467		Qy	4321	CTGTGGTTTTTTTGGCTTATGTTTCCATCAGTTTATCTGTTCACTTATAATGCTTT	4380	
Qy	3241	TTGTGATAGCTTGGCTTTTCAAGGACATGGACAAATTTTTCAGGTTTGTGTTTCTGCCCTAG	3300		Db	4548	CTGTGGTTTTTTTGGCTTATGTTTCCATCAGTTTATCTGTTCACTTATAATGCTTT	4607	
Db	3468	TTGTGATAGCTTGGCTTTTCAAGGACATGGACAAATTTTTCAGGTTTGTGTTTCTGCCCTAG	3527		Qy	4381	CTTTCACCTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCATTTATCTGTTGG	4440	
Qy	3301	ACAGTCAATTCAAATTTGGGTGTCAATTTCTTATGGTGTTCCTATGACGACTTTTGGAAAGCG	3360		Db	4608	CTTTCACCTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCATTTATCTGTTGG	4667	
Db	3528	ACAGTCAATTCAAATTTGGGTGTCAATTTCTTATGGTGTTCCTATGACGACTTTGGAAGAGCG	3587		Qy	4441	CAGCGTTGCTGTTGATGCAATCACTGAAATAACTTTTCTTTTATGGGATACAAATGTCAA	4500	
Qy	3361	TATTTTTAAAGCTTGAAGTTGAAGCAGAAATTTGACCAAGCAGATATATAGTGTATTTACTC	3420		Db	4668	CAGCGTTGCTGTTGATGCAATCACTGAAATAACTTTTCTTTATGGGATACAAATGTCAA	4727	
Db	3588	TATTTTTAAAGCTTGAAGTTGAAGCAGAAATTTGACCAAGCAGATATATAGTGTATTTACTC	3647		Qy	4501	CTATTCTTCATTTATGCCCTTTTGTATCATCATTTCCAAATCTATCCACTTTCTAGGTTGCCGA	4560	
Qy	3421	AGCAGCCACTGGAGGAGAAATGGATTCAAAATCTTTTGTATGAATGGAACAGAGCTTAC	3480		Db	4728	CTATTCTTCATTTATGCCCTTTTGTATCATCATTTCCAAATCTATCCACTTTCTAGGTTGCCGA	4787	
Db	3648	AGCAGCCACTGGAGGAGAAATGGATTCAAAATCTTTTGTATGAATGGAACAGAGCTTAC	3707		Qy	4561	TTTCTTTTCAATAAGATTTTCTTGAAGAAATGTACGAAATAATGTGGACACCTATATCCAT	4620	
Qy	3481	TTATTTCTTGAAACCAAGGCTTCTCTAGTGAGCCACCATGAGCTTTTGGAAACAAACAGA	3540		Db	4788	TTTCTTTTCAATAAGATTTTCTTGAAGAAATGTACGAAATAATGTGGACACCTATAATCCAT	4847	
Db	3708	TTATTTCTTGAAACCAAGGCTTCTCTAGTGAGCCACCATGAGCTTTTGGAAACAAACAGA	3767		Qy	4621	GGGATAGGCTTTTTCAGTAGCTGTTTATATCGCTTACTCTGCAAGTGTGCTGCTGATTTTCC	4680	
Qy	3541	TGTATACAATAGCAAGCTTTCTTTTACCTTTGAACCTGGAAGTGAATTAATTCAGTGAGAT	3600		Db	4848	GGGATAGGCTTTTTCAGTAGCTGTTTATATCGCTTACTCTGCAAGTGTGCTGCTGATTTTCC	4907	
Db	3768	TGTATACAATAGCAAGCTTTCTTTTACCTTTGAACCTGGAAGTGAATTAATTCAGTGAGAT	3827		Qy	4681	TCTTACAATPACTATGAGAAAAAATATGAGGAGATCAATAAGAAAGATCCCTTTTTC	4740	
Qy	3601	CAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTTTGGTTTCATC	3660						

```
Db 4908 TCCTTACAATACTATGAGAAAAAATATGAGGAGAGATCAATAAGAAAAATCCCTTTTCA 4967
Qy 4741 GAAACCTTTTCAAGCAAGCTCTAAAAATAGGAAGCTTCCAGAACACACAGACAAATGAGGATG 4800
Db 4968 GAAACCTTTTCAAGCAAGCTCTAAAAATAGGAAGCTTCCAGAACACACAGACAAATGAGGATG 5027
Qy 4801 AAGATGAAGATGCTCAAGCTGGAAGACTAAAGGTCAAAAGCTGATGGGTTCAGGTT 4860
Db 5028 AAGATGAAGATGCTCAAGCTGGAAGACTAAAGGTCAAAAGCTGATGGGTTCAGGTT 5087
Qy 4861 GTGAGGAGAACCAATCAATTTGGTGCAGCAATTTGCATTAAGATATGATGACAGAAAG 4920
Db 5088 GTGAGGAGAACCAATCAATTTGGTGCAGCAATTTGCATTAAGATATGATGACAGAAAG 5147
Qy 4921 ATTTCTCTTTTCAAGAAAAAGTAAAGAAAGTGGCAACTAAATACATCTCTTTCTGTGTGA 4980
Db 5148 ATTTCTCTTTTCAAGAAAAAGTAAAGAAAGTGGCAACTAAATACATCTCTTTCTGTGTGA 5207
Qy 4981 AAAAAAGAGAGATCTTTAGGACTATTTGGGTCCAAATGGTGTGGCAAAAGCACAAATTATTA 5040
Db 5208 AAAAAAGAGAGATCTTTAGGACTATTTGGGTCCAAATGGTGTGGCAAAAGCACAAATTATTA 5267
Qy 5041 ATATTCTGGTGTGTGATTTGAACCAACTTTCAGGCCAGGTATTTTAGGAGATATTTCTT 5100
Db 5268 ATATTCTGGTGTGTGATTTGAACCAACTTTCAGGCCAGGTATTTTAGGAGATATTTCTT 5327
Qy 5101 CAGAGACAACTGAAGATGATGATTTCACTGAAGTGTATGGGTCTACTGTCCTCAGATAAAC 5160
Db 5328 CAGAGACAACTGAAGATGATGATTTCACTGAAGTGTATGGGTCTACTGTCCTCAGATAAAC 5387
Qy 5161 CTTTGTGGCCAGATPACTACATTCAGGAACATTTTGAATTTTATGGAGCTGTCAAAAGGAA 5220
Db 5388 CTTTGTGGCCAGATPACTACATTCAGGAACATTTTGAATTTTATGGAGCTGTCAAAAGGAA 5447
Qy 5221 TGAGTGAAGTGCATGAAAGAGTCAATAGTGAATTAACACATGCACCTTGATTTAAAG 5280
Db 5448 TGAGTGAAGTGCATGAAAGAGTCAATAGTGAATTAACACATGCACCTTGATTTAAAG 5507
Qy 5281 AACATCTTCAGAGACTGTAAGAAACTTACTCTGAGGAATCAAAAGCAAGCTGTGTTTTCG 5340
Db 5508 AACATCTTCAGAGACTGTAAGAAACTTACTCTGAGGAATCAAAAGCAAGCTGTGTTTTCG 5567
Qy 5341 CTCTAAGTATGCTAGGGAATCCTCAGATTAATTTTGTGTAGATGAACCATCTACAGGTATGG 5400
Db 5568 CTCTAAGTATGCTAGGGAATCCTCAGATTAATTTTGTGTAGATGAACCATCTACAGGTATGG 5627
Qy 5401 ATCCAAAGCCAAACAGCAGATGTGGCAGCAATTCGAACTGCAATTTAAACACAGAAAGC 5460
Db 5628 ATCCAAAGCCAAACAGCAGATGTGGCAGCAATTCGAACTGCAATTTAAACACAGAAAGC 5687
Qy 5461 GGGCTGCTATTCTGACCACTCCTATATATGAGGAGGCGAGAGGCTGTCTGTGATCGAGTAG 5520
Db 5688 GGGCTGCTATTCTGACCACTCCTATATATGAGGAGGCGAGAGGCTGTCTGTGATCGAGTAG 5747
Qy 5521 CTATCATGGTGTCTGGGCAAGTAAAGATGATTCGAAACAGATACAACTCTAAAGAGTAAAT 5580
Db 5748 CTATCATGGTGTCTGGGCAAGTAAAGATGATTCGAAACAGATACAACTCTAAAGAGTAAAT 5807
Qy 5581 TTGAAAGAGGCTACTTTTGGAAATTTAAATTTGAAGAGTATGATGATGATGATGATGATGATG 5640
Db 5808 TTGAAAGAGGCTACTTTTGGAAATTTAAATTTGAAGAGTATGATGATGATGATGATGATGATG 5867
Qy 5641 ACCGCCCTCAAGAGAAATTCAGTATATTTTCCCAATTCAGCCGTCAGGAAAGTTTTT 5700
Db 5868 ACCGCCCTCAAGAGAAATTCAGTATATTTTCCCAATTCAGCCGTCAGGAAAGTTTTT 5927
Qy 5701 CTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTTCACAATCTTTTT 5760
Db 5928 CTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTTCACAATCTTTTT 5987
Qy 5761 TTAAGCTGGAAGAGCTAAACATGCTTTTGCATTTGAAGAAATATAGCTTTTCTCAAGCAA 5820
Db 5988 TTAAGCTGGAAGAGCTAAACATGCTTTTGCATTTGAAGAAATATAGCTTTTCTCAAGCAA 6047
```

```
Qy 5821 CATTTGAAACAGGTTTTTTGTAGAACTCATAAGAAACAAGAGGAGGAATAAATAGTTGTG 5880
Db 6048 CATTTGAAACAGGTTTTTTGTAGAACTCATAAGAAACAAGAGGAGGAATAAATAGTTGTG 6107
Qy 5881 GAACCTTTAAACAGCACACTTTTGGTGGAAAGCAACACAGAAAGATAGAGTAGTATTTTGAA 5940
Db 6108 GAACCTTTAAACAGCACACTTTTGGTGGAAAGCAACACAGAAAGATAGAGTAGTATTTTGAA 6167
Qy 5941 TTTGTATTGTCGGTCTGCTTACTCTGGGACTCTCTTTCTTTTCACTTAATTTTAACTTTGG 6000
Db 6168 TTTGTATTGTCGGTCTGCTTACTCTGGGACTCTCTTTCTTTTCACTTAATTTTAACTTTGG 6227
Qy 6001 TTTAAAAAGTTTTTTTATTTGGAATGTTAACTGGAGAACCAAGAACCGCACCTTGAATTTTTC 6060
Db 6228 TTTAAAAAGTTTTTTTATTTGGAATGTTAACTGGAGAACCAAGAACCGCACCTTGAATTTTTC 6287
Qy 6061 TAACTCTCTTAATTTGAATGCTGTGTTGTGTTGTTTCTTTTCTTTTAAATAAACGTAT 6120
Db 6288 TAACTCTCTTAATTTGAATGCTGTGTTGTGTTTCTTTTCTTTTAAATAAACGTAT 6347
Qy 6121 GTATAATTAAGTGAA 6135
Db 6348 GTATAATTAAGTGAA 6362

RESULT 4
AX537473
LOCUS AX537473
DEFINITION Sequence 4 from Patent WO02070690.
ACCESSION AX537473
VERSION AX537473.1 GI:25269282
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chen, H., Kilinski, L. and le Bihan, S.
TITLE Abca5 transporter and uses thereof
JOURNAL Patent: WO 02070690-A 4 12-SEP-2002;
Active Pass Pharmaceuticals, Inc. (CA)
FEATURES
source
1. .5475
Location/Qualifiers
/mol_type="genomic DNA"
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1705 a 917 c 1048 g 1804 t 1 others
ORIGIN
Query Match 82.3%; Score 5373.2; DB 6; Length 5475;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5375; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 994 AGGTTTATTCAGAAAAACATGTCACCTGCATTTAGGAGGATAGGAGTTTGGAGACAGACCA 1053
Db 97 AGGTTTATTCAGAAAAACATGTCACCTGCATTTAGGAGGATAGGAGTTTGGAGACAGACCA 156
Qy 1054 GAACCTCTTACTGAAGAATTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1113
Db 157 GAACCTCTTACTGAAGAATTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 216
Qy 1114 AAATCTTTTCCACTATTTTCTTTTATTTTGGTTTATTAATTAATTAATTAATTAATTAATTAAT 1173
Db 217 AAATCTTTTCCACTATTTTCTTTTATTTTGGTTTATTAATTAATTAATTAATTAATTAATTAAT 276
Qy 1174 ATAAGAAATATGAAGAAGTGCCTTAATATAGAACTCAATCTATGGACAAAGTTTACTCTTT 1233
Db 277 ATAAGAAATATGAAGAAGTGCCTTAATATAGAACTCAATCTATGGACAAAGTTTACTCTTT 336
Qy 1234 CTAACTTAATTTCTTGGATATFATCCAGTGATTAATTAATTAATTAATTAATTAATTAATTAAT 1293
Db 337 CTAACTTAATTTCTTGGATATFATCCAGTGATTAATTAATTAATTAATTAATTAATTAATTAAT 396
```

QY 1294 TGTCTACTGATCATCTACCTGATGTGCATTAATTAAGAAATATACAAATGAAAGAAA 1353  
DB 397 TGTCTACTGATCATCTACCTGATGTGCATTAATTAAGAAATATACAAATGAAAGAAA 456  
QY 1354 TGTAAACATCCAGTCTCTCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAAGACTCCA 1413  
DB 457 TGTAAACATCCAGTCTCTCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAAGACTCCA 516  
QY 1414 TGTCTCTATGAACCTTGGTTTTTTTCCGTGAATGAATTCACAGTATCTTCTATATATATGATTT 1473  
DB 517 TGTCTCTATGAACCTTGGTTTTTTTCCGTGAATGAATTCACAGTATCTTCTATATATATGATTT 576  
QY 1474 CAAGAGCTGGCTCTCAAAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTTCACAG 1533  
DB 577 CAAGAGCTGGCTCTCAAAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTTCACAG 636  
QY 1534 TTTTACAAGCATCCATAGATGCTGCCATTTATACAGTTGAAGCAATGTTTCTCTTTGGA 1593  
DB 637 TTTTACAAGCATCCATAGATGCTGCCATTTATACAGTTGAAGCAATGTTTCTCTTTGGA 696  
QY 1594 AGGAGCTGGAGTCAACTAAAGCTGTTATTAAGGAGAACTGCTGTTGAGAAATAGATA 1653  
DB 697 AGGAGCTGGAGTCAACTAAAGCTGTTATTAAGGAGAACTGCTGTTGAGAAATAGATA 756  
QY 1654 CCTTTCCCGAGGAGTAATTTTAAATATACCTAGTATAGCATTTTCCACCTTTTGGATACT 1713  
DB 757 CCTTTCCCGAGGAGTAATTTTAAATATACCTAGTATAGCATTTTCCACCTTTTGGATACT 816  
QY 1714 TTTTGGCAATTCATATCGTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGATAA 1773  
DB 817 TTTTGGCAATTCATATCGTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGATAA 876  
QY 1774 TGGGACTTCATATGATACCTTTGGCTTTTCCTGGGTTCTTCTATATACAAGTTTAAATTT 1833  
DB 877 TGGGACTTCATATGATACCTTTGGCTTTTCCTGGGTTCTTCTATATACAAGTTTAAATTT 936  
QY 1834 TTCTATGTCCCTCTTATGGGAGTCAATGGGACAGCTCTTGTATTTTCCCTCAAGTA 1893  
DB 937 TTCTATGTCCCTCTTATGGGAGTCAATGGGACAGCTCTTGTATTTTCCCTCAAGTA 996  
QY 1894 GCAGCATTTGATATTTCTCTTTTTCCTTTATGGATATCATCTGATTTTTTGGCTT 1953  
DB 997 GCAGCATTTGATATTTCTCTTTTTCCTTTATGGATATCATCTGATTTTTTGGCTT 1056  
QY 1954 TAATGCTGACACCTCTTTTAAAGATCAAAACATGTGGGAATAGTTGAAATTTTGTGTA 2013  
DB 1057 TAATGCTGACACCTCTTTTAAAGATCAAAACATGTGGGAATAGTTGAAATTTTGTGTA 1116  
QY 2014 CTGTGGCTTTTGGATTTATTTGGCTTTATGATATATCCTCATAGAAAGTTTTTCCAAATCGT 2073  
DB 1117 CTGTGGCTTTTGGATTTATTTGGCTTTATGATATATCCTCATAGAAAGTTTTTCCAAATCGT 1176  
QY 2074 TAGTGGCTTTTTCAGTCTCTTCTGTCACCTGATCTTTTGTGATGTTGATGACAGGTCA 2133  
DB 1177 TAGTGGCTTTTTCAGTCTCTTCTGTCACCTGATCTTTTGTGATGTTGATGACAGGTCA 1236  
QY 2134 TGCATTTAGAAAGATTTTAAAGAGTGTCTTCAATTTCAAAATTTGACTGCAAGGCCCATATC 2193  
DB 1237 TGCATTTAGAAAGATTTTAAAGAGTGTCTTCAATTTTCAAAATTTGACTGCAAGGCCCATATC 1296  
QY 2194 CTCTATATTTACAAATTTATCATGCTCACCTTAATAGTATATTTCTATGCTCTTGGCTG 2253  
DB 1297 CTCTATATTTACAAATTTATCATGCTCACCTTAATAGTATATTTCTATGCTCTTGGCTG 1356  
QY 2254 TCTATCTTTGATCAAGTCAATTTCCAGGGGAAATTTGGCTTTACGGAGATCATCTTTATATTTTC 2313  
DB 1357 TCTATCTTTGATCAAGTCAATTTCCAGGGGAAATTTGGCTTTACGGAGATCATCTTTATATTTTC 1416  
QY 2314 TGAAGCCTTCATATTTGCTCAAGAGTAAAGAAATTTATGAGGAGTTATCAGAGGGCAATG 2373  
DB 1417 TGAAGCCTTCATATTTGCTCAAGAGCAAGAAATTTATGAGGAGTTATCAGAGGGCAATG 1476

QY 2374 TTAATGGAAATATTAGTTTTAGTGAAATTTATGAGCCAGTTTCTTTCAGAAATTTGTAGAA 2433  
DB 1477 TTAATGGAAATATTAGTTTTAGTGAAATTTATGAGCCAGTTTCTTTCAGAAATTTGTAGAA 1536  
QY 2434 AAGAAAGCCATTAAGAAATTTAGTGGTATTAGAAAGATATCAGAAAGAGAGGTGAAATGTGG 2493  
DB 1537 AAGAAAGCCATTAAGAAATTTAGTGGTATTAGAAAGATATCAGAAAGAGAGGTGAAATGTGG 1596  
QY 2494 AGGCTTTTCAGAAATTTTGTTCATTTGACATATATGAGGTCAGATTACTGCCTTTACTTTGCC 2553  
DB 1597 AGGCTTTTCAGAAATTTTGTTCATTTGACATATATGAGGTCAGATTACTGCCTTTACTTTGCC 1656  
QY 2554 ACAGTGGAAACAGGAAAGATACATTTGATGAATATTTCTTTTGGGACTCTGCCACCTCTG 2613  
DB 1657 ACAGTGGAAACAGGAAAGATACATTTGATGAATATTTCTTTTGGGACTCTGCCACCTCTG 1716  
QY 2614 ATGGGTTTTGCATCTATATATATATGGAACACAGAGTCTCAGAAATATAGTAAATGTTTGAAGCA 2673  
DB 1717 ATGGGTTTTGCATCTATATATATGGAACACAGAGTCTCAGAAATATAGTAAATGTTTGAAGCA 1776  
QY 2674 GAAAATGATTTGGCAATTTTGTCCACAGTTAGATATACATTTTGTATGTTTTTGACAGTAGAG 2733  
DB 1777 GAAAATGATTTGGCAATTTTGTCCACAGTTAGATATACATTTTGTATGTTTTTGACAGTAGAG 1836  
QY 2734 AAAATTTTCAATTTTGGCTTTCAATCAAAGGATACCAAGCCAAACAATATAATACAAAG 2793  
DB 1837 AAAATTTTCAATTTTGGCTTTCAATCAAAGGATACCAAGCCAAACAATATAATACAAAG 1896  
QY 2794 TGCAGAGGTTTTACTAGATTTAGACATGCGAGCTATCAAAGATACCAAGCTTAAAAAT 2853  
DB 1897 TGCAGAGGTTTTTACTAGATTTAGACATGCGAGCTATCAAAGATACCAAGCTTAAAAAT 1956  
QY 2854 TAAAGTGGTGGTCAAAAAAGAAAGCTGTCTATTAGGAATTTGCTTCTTGGGAACCCAAAGA 2913  
DB 1957 TAAAGTGGTGGTCAAAAAAGAAAGCTGTCTATTAGGAATTTGCTTCTTGGGAACCCAAAGA 2016  
QY 2914 TACTGCTGCTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATTTGATGGA 2973  
DB 2017 TACTGCTGCTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATTTGATGGA 2076  
QY 2974 ATCTTTTAAAAATACAGAAAAAGCCAAATCGGGTGAACAGTGTTCAGTACTCATTTTATGAGATG 3033  
DB 2077 ATCTTTTAAAAATACAGAAAAAGCCAAATCGGGTGAACAGTGTTCAGTACTCATTTTATGAGATG 2136  
QY 3034 BAGCTGACATTTCTGACATAGGAAAGCTGTGATATCAAGGAATGCTGAAATGTGTTG 3093  
DB 2137 BAGCTGACATTTCTGACATAGGAAAGCTGTGATATCAAGGAATGCTGAAATGTGTTG 2196  
QY 3094 GTTCTTTCAATGTTCTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 3153  
DB 2197 GTTCTTTCAATGTTCTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2256  
QY 3154 ACAAAATATTTGGCCACAGAAATCTCTTTTCACTGGTTTAAACAACATATACCTGAGCTA 3213  
DB 2257 ACAAAATATTTGGCCACAGAAATCTCTTTTCACTGGTTTAAACAACATATACCTGAGCTA 2316  
QY 3214 CTTTATTTACACAGAAATGACCAACAACTTGTGTATAGCTTGCCTTTCAAGGACATGGA 3273  
DB 2317 CTTTATTTACACAGAAATGACCAACAACTTGTGTATAGCTTGCCTTTCAAGGACATGGA 2376  
QY 3274 AATTTTTCAGGTTTGTGTTTTCTGCCCTTAGACAGTCAATTTCAAAATTTGGGTGTCTTTATG 3333  
DB 2377 AATTTTTCAGGTTTGTGTTTTCTGCCCTTAGACAGTCAATTTCAAAATTTGGGTGTCTTTATG 2436  
QY 3334 GTGTTTTCCATGACGATTTTGAAGACGATTTTAAAGCTAGAGTTAGAGCAGAAATG 3393  
DB 2437 GTGTTTTCCATGACGATTTTGAAGACGATTTTAAAGCTAGAGTTAGAGCAGAAATG 2496  
QY 3394 ACCAAGCAGATTTATAGTGTATTTACTCAGAGCCACTGAGGAGAAATGGAATTCAAAT 3453  
DB 2497 ACCAAGCAGATTTATAGTGTATTTACTCAGAGCCACTGAGGAGAAATGGAATTCAAAT 2556  
QY 3454 CTTTTGATGAATGGAACAGAGCTTTACTTTTCTTCTTCTGAAACCAAGGCTTCTCTAGTGA 3513

|||||  
2557 CTTTGTGAATGGAAACAGAGCTTACTTATCTTTCTGAAACCAAGGCTTCTTAGTGA 2616  
QY  
3514 GCACCATGAGCCTTTTGAACAACAAGATGATACAAATAGCAAAAGTTTCATTCTCTTACCT 3573  
Db  
2617 GCACCATGAGCCTTTGAAACAACAAGATGATACAAATAGCAAAAGTTTCATTCTTACCT 2676  
QY  
3574 TGAACGCTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTCACAG 3633  
Db  
2677 TGAACGCTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTCACAG 2736  
QY  
3634 TTGAGATTTTATGTTTTGTTGTTTCATCCTCTTTTAAATATGCTGTGGTTCCCATCAAAC 3693  
Db  
2737 TTGAGATTTTATGTTTTGTTGTTTCATCCTCTTTTAAATATGCTGTGGTTCCCATCAAAC 2796  
QY  
3694 TTGTTCCAGACTTATATTTCTTAAACCTGAGACAAACACATAAATACAAACAAAGTC 3753  
Db  
2797 TTGTTCCAGACTTATATTTCTTAAACCTGAGACAAACACATAAATACAAACAAAGTC 2856  
QY  
3754 TGCTTCTTCAAAATTCGTGACTCAGATATCAGTGATCTTATTAGCTTTTTTTCACAAAGCC 3813  
Db  
2857 TGCTTCTTCAAAATTCGTGACTCAGATATCAGTGATCTTATTAGCTTTTTTTCACAAAGCC 2916  
QY  
3814 AGAACATAATGCTGACGATGATTAATGACAGTGACTATGATCCGTGGCTCCCATAGTG 3873  
Db  
2917 AGAACATAATGCTGACGATGATTAATGACAGTGACTATGATCCGTGGCTCCCATAGTG 2976  
QY  
3874 CGGCTTTAAATGATGCATTCAGAAAGGACTATGTTTTGAGCTGTTTTTCAACAGTA 3933  
Db  
2977 CGGCTTTAAATGATGCATTCAGAAAGGACTATGTTTTGAGCTGTTTTTCAACAGTA 3036  
QY  
3934 CTATGGTTTATCTTTTACCTATATTTAGTGAATATCATTTAGTAAGTACTATCTTTTATCATTT 3993  
Db  
3037 CTATGGTTTATCTTTTACCTATATTTAGTGAATATCATTTAGTAAGTACTATCTTTTATCATTT 3096  
QY  
3994 TAAATGTCAGTGAACCAATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAAGTGA 4053  
Db  
3097 TAAATGTCAGTGAACCAATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAAGTGA 3156  
QY  
4054 TAGTTTTTAAATGAGCTGATTTTCAAGCAGCTTTGCTTGGATCAATTTTACTGCAA 4113  
Db  
3157 TAGTTTTTAAATGAGCTGATTTTCAAGCAGCTTTGCTTGGATCAATTTTACTGCAA 3216  
QY  
4114 TGCCACCTTACTTTGCCATGGAATGCAGAGATCATAGATCAAAAGCTTTACTCAAC 4173  
Db  
3217 TGCCACCTTACTTTGCCATGGAATGCAGAGATCATAGATCAAAAGCTTTACTCAAC 3276  
QY  
4174 TTAACTTTTCAGGCTTTTGGCATCTGCATATTTGGAATGGACAGCTGTTGTTGATACC 4233  
Db  
3277 TTAACTTTTCAGGCTTTTGGCATCTGCATATTTGGAATGGACAGCTGTTGTTGATACC 3336  
QY  
4234 CCTTATTTTATCATTTCTTATTTTGTGATGCTAGGAAGCTTACTGGCATTTTCATTATGGAT 4293  
Db  
3337 CCTTATTTTATCATTTCTTATTTTGTGATGCTAGGAAGCTTATTTGGCATTTTCATTATGGAT 3396  
QY  
4294 TATATTTTATCTGTAAGTTCTTCTGCTGGTTTTTTTGGCTTATTTGGTTTATGTTCCAT 4353  
Db  
3397 TATATTTTATCTGTAAGTTCTTCTGCTGGTTTTTTTGGCTTATTTGGTTTATGTTCCAT 3456  
QY  
4354 CAGTATTTCTGTCATTTATATGCTTTCTTTTCACTTTTAAAGAAATTTTAAATACCAAAG 4413  
Db  
3457 CAGTATTTCTGTCATTTATATGCTTTCTTTTCACTTTTAAAGAAATTTTAAATACCAAAG 3516  
QY  
4414 AATTTTGGTCATTTATCTATCTGTTGGAGGTTGCTGATGTAATGCAATCACTGAAATAA 4473  
Db  
3517 AATTTTGGTCATTTATCTATTTCTGTTGGAGGTTTGGCTTGTATTTGCAATCACTGAAATAA 3576  
QY  
4474 CTTTCTTTTATGGGATACAAATTCGAATTTCTTCAATATGCTTTTGTATCATCATTC 4533  
Db  
3577 CTTTCTTTTATGGGATACAAATTCGAATTTCTTCAATATGCTTTTGTATCATCATTC 3636  
QY  
4534 CAATCTATCCACTTCTAGGTTGCCGTGATTTCTTTTCAATAAGATTTCTTGGAAAGATGTAC 4593  
Db

Db  
3637 CAATCTATCCACTTCTAGTGTGCTGATTTCTTTCAATAAGATTTCTTGGAAAGATGTAC 3696  
QY  
4594 GAAAAATGTGGACACCTATAATCCATGGATAGGCTTTTCAGTAGCTGTTTATATCGCCTT 4653  
Db  
3697 GAAAAATGTGGACACCTATAATCCATGGATAGGCTTTTCAGTAGCTGTTTATATCGCCTT 3756  
QY  
4654 ACCTGAGTGTGATGTTGATTTTCTCTTACAATCTATGAGAAAAAATATGAGGCA 4713  
Db  
3757 ACCTGAGTGTGATGTTGATTTTCTCTTACAATCTATGAGAAAAAATATGAGGCA 3816  
QY  
4714 GATCAATAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAAGAAAGTCTTAAAAATAGGAGC 4773  
Db  
3817 GATCAATAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAAGAAAGTCTTAAAAATAGGAGC 3876  
QY  
4774 TTCCAGAACCAACAGACCAATGAGATGAAGATGAAGATGTCAAGCTGAAAGACTAAAGG 4833  
Db  
3877 TTCCAGAACCAACAGACCAATGAGATGAAGATGAAGATGTCAAGCTGAAAGACTAAAGG 3936  
QY  
4834 TCAAGAGCTGATGGGTGGCTGAGTGTGAGAGAAACCATCCATTTATGGTCAGCAATTT 4893  
Db  
3937 TCAAGAGCTGATGGGTGGCTGAGTGTGAGAGAAACCATCCATTTATGGTCAGCAATTT 3996  
QY  
4894 TGCATAAAGATATGATGACAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGG 4953  
Db  
3997 TGCATAAAGATATGATGACAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGG 4056  
QY  
4954 CAACTAAATACATCTCTTTCTGTGTGAAAAAGAGAGATCTTAGGACTATTGGGTCCAA 5013  
Db  
4057 CAACTAAATACATCTCTTTCTGTGTGAAAAAGAGAGATCTTAGGACTATTGGGTCCAA 4116  
QY  
5014 ATGGTCTGTGCAAAAGCACAAATTTAATAATTTCTGGTTGGTGATATTGAACCAACTTCAG 5073  
Db  
4117 ATGGTCTGTGCAAAAGCACAAATTTAATAATTTCTGGTTGGTGATATTGAACCAACTTCAG 4176  
QY  
5074 GCAGAGTATTTTATGAGAGATTTCTTTCAGAGACAAAGTGAAGATGATGATTCATGAAT 5133  
Db  
4177 GCAGAGTATTTTATGAGAGATTTCTTTCAGAGACAAAGTGAAGATGATGATTCATGAAT 4236  
QY  
5134 GTATGGTTACTGCTCTCAGATAAACCCCTTTGTGGCCAGATCTACATTCAGAGAAACATTT 5193  
Db  
4237 GTATGGTTACTGCTCTCAGATAAACCCCTTTGTGGCCAGATCTACATTCAGAGAAACATTT 4296  
QY  
5194 TTGAAATTTTATGAGAGCTGTCAAAAGGAATGAGTCAAGTGAACATGAAGAAAGTCAATAGTC 5253  
Db  
4297 TTGAAATTTTATGAGAGCTGTCAAAAGGAATGAGTCAAGTGAACATGAAGAAAGTCAATAGTC 4356  
QY  
5254 GAATAACATGACATTTGATTTTAAAGAACATCTTTCAGAGACTGTAAAGAAAGTCACTG 5313  
Db  
4357 GAATAACATGACATTTGATTTTAAAGAACATCTTTCAGAGACTGTAAAGAAAGTCACTG 4416  
QY  
5314 CAGGAATCAAAACGAAGTTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTAATTT 5373  
Db  
4417 CAGGAATCAAAACGAAGTTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTAATTT 4476  
QY  
5374 TGCTAGATGAACCATCTACAGGTATGGATCCCAAGCCAAACAGCACATGTGGCGAGCAA 5433  
Db  
4477 TGCTAGATGAACCATCTACAGGTATGGATCCCAAGCCAAACAGCACATGTGGCGAGCAA 4536  
QY  
5434 TTGAACTGTCAATTTAAACAGAAAGCGGCTGCTATTTCTGACCATCTCCTATATGAGG 5493  
Db  
4537 TTGAACTGTCAATTTAAACAGAAAGCGGCTGCTATTTCTGACCATCTCCTATATGAGG 4596  
QY  
5494 AGCAGAGGCTGTCTGTATCGAGTAGCTATCATGTTGCTGGGAGTAAAGATGATTCG 5553  
Db  
4597 AGCAGAGGCTGTCTGTATCGAGTAGCTATCATGTTGCTGGGAGTAAAGATGATTCG 4656  
QY  
5554 GAACAGTCAACATCTTAAAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTTAAATTTGA 5613  
Db  
4657 GAACAGTCAACATCTTAAAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTTAAATTTGA 4716  
QY  
5614 AGGACTGGATGAGAAACCTTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCC 5673  
Db  
4717 AGGACTGGATGAGAAACCTTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCC 4776



```
QY 5674 CAAATGCAAGCGTCAGGAAAGTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAG 5733
Db 4777 CAAATGCAAGCGTCAGGAAAGTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAG 4836
QY 5734 ATGTTAGTCCCTTTCACAACTCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGCCA 5793
Db 4837 ATGTTAGTCCCTTTCACAACTCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGCCA 4896
QY 5794 TTGNAGATATAGCTTTTCTCAAGCACATTTGGAAGAGCTTTTGTAGAACTCACTAAAG 5853
Db 4897 TTGAAGAATATAGCTTTTCTCAAGCAACATTTGGAAGAGCTTTTGTAGAACTCACTAAAG 4956
QY 5854 AACAAAGAGGAGGAGATAATAGTTGTGGAACCTTTTAAACAGCACACTTTTGTGGGAACGAA 5913
Db 4957 AACAAAGAGGAGGAGATAATAGTTGTGGAACCTTTTAAACAGCACACTTTTGTGGGAACGAA 5016
QY 5914 CACAAGAGATAGAGTAGTATTTTGAATTTGTATTTGTTCGGTCTGCTTACTGGGACTTCT 5973
Db 5017 CACAAGAGATAGAGTAGTATTTTGAATTTGTATTTGTTCGGTCTGCTTACTGGGACTTCT 5076
QY 5974 TTCTTTTCACTTAATTTTAACTTTTGTGTTTAAAGTTTATTTTGAATTTGTGGAATGGAATGGA 6033
Db 5077 TTCTTTTCACTTAATTTTAACTTTTGTGTTTAAAGTTTATTTTGAATTTGTGGAATGGAATGGA 5136
QY 6034 GAACCAAGAGCGCACTTTGAAATTTTCTTAAGCTCTTAAATTTGAAATGCTGCTGTTGTG 6093
Db 5137 GAACCAAGAGCGCACTTTGAAATTTTCTTAAGCTCTTAAATTTGAAATGCTGCTGTTGTG 5196
QY 6094 TTTTGTCTTTCTTTTAAATAAAACGATATGATATAATTAAGTGAAGCTGCAATGTTTGTATGA 6153
Db 5197 TTTTGTCTTTCTTTTAAATAAAACGATATGATATAATTAAGTGAAGCTGCAATGTTTGTATGA 5256
QY 6154 AGTATATGAACTATATAGTTTGTATGTCATCTTTTTCACCATTGAGAAAGCTGCTTCT 6213
Db 5257 AGTATATGAACTATATAGTTTGTATGTCATCTTTTTCACCATTGAGAAAGCTGCTTCT 5316
QY 6214 GAATTTGTGATTTTAAAGGAATTTGTAATAGAAATAGTTTATTTTAAAGTATCTTTAAAGTT 6273
Db 5317 GAATTTGTGATTTTAAAGGAATTTGTAATAGAAATAGTTTATTTTAAAGTATCTTTAAAGTT 5376
QY 6274 TATGCCATCTTCTTTAAATAAGTAGTAATGTTCCAATCTAAATTAATAAAACCTAATACATAA 6333
Db 5377 TATGCCATCTTCTTTAAATAAGTAGTAATGTTCCAATCTAAATTAATAAAACCTAATACATAA 5436
QY 6334 CTAATGATAGAAAGATACATAAAGCAATGTGAAGTT 6372
Db 5437 CTAATGATAGAAAGATACATAAAGCAATGTGAAGTT 5475

RESULT 5
AX537470 LOCUS 5463 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 1 from Patent W02070690.
ACCESSION AX537470
VERSION AX537470.1 GI:25269277
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chen, H., Kilinski, L. and le Bihan, S.
TITLE Abca5 transporter and uses thereof
JOURNAL Patent: WO 02070690-A 1 12-SEP-2002;
Active Pass Pharmaceuticals, Inc. (CA)
FEATURES
source
1. .5463
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 1699 a 912 c 1055 g 1794 t 3 others
ORIGIN
```

```
Query Match 80.5%; Score 5255.8; DB 6; Length 5463;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 5340; Conservative 0; Mismatches 8; Indels 50; Gaps 3;
QY 994 AGGTTTATTCAGAAAACATGTCCTCACTGCAATTTAGGGAGGTAGGAGTTTGGAGACAGACCA 1053
Db 97 AGGTTTATTCAGAAAACATGTCCTCACTGCAATTTAGGGAGGTAGGAGTTTGGAGACAGACCA 156
QY 1054 GAACACTTCTACTGAGAAATTTACTTTAAATTAATGAGAACCAAAAGAGTAGTGTTCAGG 1113
Db 157 GAACACTTCTACTGAGAAATTTACTTTAAATTAATGAGAACCAAAAGAGTAGTGTTCAGG 216
QY 1114 AAATTTCTTTTCCACTATTTTATTTTGGTTAAATTAATTAAGCATGATGCAATCCAA 1173
Db 217 AAATTTCTTTTCCACTATTTTATTTTGGTTAAATTAATTAAGCATGATGCAATCCAA 276
QY 1174 ATAAGAAATATGAAGAAGTGCTTAATATAGAACTCAATCCCTATGAGCAAGTTTACTCTTT 1233
Db 277 ATAAGAAATATGAAGAAGTGCTTAATATAGAACTCAATCCCTATGAGCAAGTTTACTCTTT 336
QY 1234 CTAATCTAATTTCTTGATATATCTCCAGTGACTTAATTAATTAAGCAAGCATGCAAGAAAG 1293
Db 337 CTAATCTAATTTCTTGATATATCTCCAGTGACTTAATTAATTAAGCAAGCATGCAAGAAAG 396
QY 1294 TGTCTACTGATCATCTACTGATGTCTAATTTACTGAAAGAAATATACAAATGAAAAAGAAA 1353
Db 397 TGTCTACTGATCATCTACTGATGTCTAATTTACTGAAAGAAATATACAAATGAAAAAGAAA 456
QY 1354 TGTTAACTCCAGTCTCTCTAAGCCGAGCAACTTTGTAGGTGTGTTTTCAAGACTCCA 1413
Db 457 TGTTAACTCCAGTCTCTCTAAGCCGAGCAACTTTGTAGGTGTGTTTTCAAGACTCCA 516
QY 1414 TGTCTATGAACTCTGTTTTTCTGATATGATTTCCAGTATCTTCTATTTATATATGGAAT 1473
Db 517 TGTCTATGAACTCTGTTTTTCTGATATGATTTCCAGTATCTTCTATTTATATATGGAAT 576
QY 1474 CAAAGAGCTGGCTGTTTCAAAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGGTTTTCACAG 1533
Db 577 CAAAGAGCTGGCTGTTTCAAAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGGTTTTCACAG 636
QY 1534 TTTTCAAGAGATCCATATAGATGTCCTATATACAGTTGAAGCAATGTTTCTCTTTGGA 1593
Db 637 TTTTCAAGAGATCCATATAGATGTCCTATATACAGTTGAAGCAATGTTTCTCTTTGGA 696
QY 1594 AGGAGCTGGCTCAACTAAAGCTGTTATATGGGAGAACTGCTGCTGTAGAAATAGATA 1653
Db 697 AGGAGCTGGCTCAACTAAAGCTGTTATATGGGAGAACTGCTGCTGTAGAAATAGATA 756
QY 1654 CCTTTCCCGAGGAGTAAATTTTAAATATACCTAGTTATATAGCAATTTTCACTTTTGGATACT 1713
Db 757 CCTTTCCCGAGGAGTAAATTTTAAATATACCTAGTTATATAGCAATTTTCACTTTTGGATACT 816
QY 1714 TTTTGGCAATTCATATCGTAGCAAAAAGAAAAAATAAAGAAATTTTAAAGATAA 1773
Db 817 TTTTGGCAATTCATATCGTAGCAAAAAGAAAAAATAAAGAAATTTTAAAGATAA 876
QY 1774 TGGGACTTCATGATACTGCTTTTGGCTTTCTGCTGTTCTTCTATATACAGTTTAAATTT 1833
Db 877 TGGGACTTCATGATACTGCTTTTGGCTTTCTGCTGTTCTTCTATATACAGTTTAAATTT 936
QY 1834 TTCTTATGTCCCTTCTATGCGAGTCAATTTGCGACAGCTTCTTTGTTATTTCTCAAGATTA 1893
Db 937 TTCTTATGTCCCTTCTATGCGAGTCAATTTGCGACAGCTTCTTTGTTATTTCTCTCAAGATTA 996
QY 1894 GCAGCATTTGATATTTCTGCTTTTTTTTCTTTATGGAATATCATCTGATTTTTTGTCTT 1953
Db 997 GCAGCATTTGATATTTCTGCTTTTTTTTCTTTATGGAATATCATCTGATTTTTTGTCTT 1056
QY 1954 TAATGCTGACACCTCTTTTAAATAAATCAAAACATGCGGAATAGTTGAAATTTTGTGTA 2013
Db 1057 TAATGCTGACACCTCTTTTAAATAAATCAAAACATGCGGAATAGTTGAAATTTTGTGTA 1116
```



QY 2014 CTGTGGCTTTTGGATTATTTGGCTTTATGATAATTCCTCATAGAAAGTTTTCCAAATCGT 2073  
Db 1117 CTGTGGCTTTTGGATTATTTGGCTTTATGATAATTCCTCATAGAAAGTTTTCCAAATCGT 1176  
QY 2074 TAGTGTGGCTTTTTCAGTCCCTTTCTGTACACGTACTTTTGTGATTTGGTATGACAGGTCA 2133  
Db 1177 TAGTGTGGCTTTTTCAGTCCCTTTCTGTACACGTACTTTTGTGATTTGGTATGACAGGTCA 1236  
QY 2134 TGCATTTAGAAAGATTTTAAATGAAAGGTGCTTCATTTTCAAAATTTTGACTGCAAGGCCATATC 2193  
Db 1237 TGCATTTAGAAAGATTTTAAATGAAAGGTGCTTCATTTTCAAAATTTTGACTGCAAGGCCATATC 1296  
QY 2194 CTCTAATATATACAAATATACATGCTCACCTTAATAGTATATTTCTATGTCCTCTGTGGCTG 2253  
Db 1297 CTCTAATATATACAAATATACATGCTCACCTTAATAGTATATTTCTATGTCCTCTGTGGCTG 1356  
QY 2254 TCTATCTTTGATCAAGTCAATTCAGGGGAATTTGGCTTTACGGAGATCATCTTTATATTTTC 2313  
Db 1357 TCTATCTTTGATCAAGTCAATTCAGGGGAATTTGGCTTTACGGAGATCATCTTTATATTTTC 1416  
QY 2314 TGAAGCCTTCATATTTGGTCAAGAGTAAAGAAATTTATGAGGAGTTATCAGAGGCAATG 2373  
Db 1417 TGAAGCCTTCATATTTGGTCAAGAGTAAAGAAATTTATGAGGAGTTATCAGAGGCAATG 1476  
QY 2374 TTAATGGAATATATAGTTTGTAGTGAATTTATGAGCCAGTTTCTTCAGAAATTTGTAGGAA 2433  
Db 1477 TTAATGGAATATATAGTTTGTAGTGAATTTATGAGCCAGTTTCTTCAGAAATTTGTAGGAA 1536  
QY 2434 AAGAGCCATAAGAAATTTAGTGGTATTCAGAAAGCATACAGAAAGAGGGTGAATTTGTG 2493  
Db 1537 AAGAGCCATAAGAAATTTAGTGGTATTCAGAAAGCATACAGAAAGAGGGTGAATTTGTG 1596  
QY 2494 AGCTTTTGAGAAATTTGTCAATTTGACATATATGAGGGTCAGATTACTGCCCTTACTTTGGCC 2553  
Db 1597 AGCTTTTGAGAAATTTGTCAATTTGACATATATGAGGGTCAGATTACTGCCCTTACTTTGGCC 1656  
QY 2554 ACAGTGGAAACAGAAAGAGTACATTTGATGAATATTTCTTTGTGGACTCTGCCACCTTTCTG 2613  
Db 1657 ACAGTGGAAACAGAAAGAGTACATTTGATGAATATTTCTTTGTGGACTCTGCCACCTTTCTG 1716  
QY 2614 ATGGGTTTGCAATCTATATATGGAACAGAGTCTCAGAAATAGATGAATTTTGAAGCAA 2673  
Db 1717 ATGGGTTTGCAATCTATATATGGAACAGAGTCTCAGAAATAGATGAATTTTGAAGCAA 1776  
QY 2674 GAAAAATGATTTGGATTGTCACAGTTAGATATACACTTTGATTTGTTTGCACAGTAGAAG 2733  
Db 1777 GAAAAATGATTTGGATTGTCACAGTTAGATATACACTTTGATTTGTTTGCACAGTAGAAG 1836  
QY 2734 AAAATTTATCAATTTTGGCTTCAATCAAGGGATACCCAGCCAAATATATATACAGAAAG 2793  
Db 1837 AAAATTTATCAATTTTGGCTTCAATCAAGGGATACCCAGCCAAATATATATACAGAAAG 1896  
QY 2794 TGCAGAAGTTTTTACTAGATTTAGACATGACAGTCTCAAGATTAACCAAGCTAAAAAT 2853  
Db 1897 TGCAGAAGTTTTTACTAGATTTAGACATGACAGTCTCAAGATTAACCAAGCTAAAAAT 1956  
QY 2854 TAAAGTGTGTGCAAAAAGAAAGCTGTCATTAGGAATTTGCTTTTGGAAACCCAAAGA 2913  
Db 1957 TAAAGTGTGTGCAAAAAGAAAGCTGTCATTAGGAATTTGCTTTTGGAAACCCAAAGA 2016  
QY 2914 TACTGTGCTAGATGAACCAACAGCTGGAAATGGACCCCTGTTCTCGACATATTTGTATGGA 2973  
Db 2017 TACTGTGCTAGATGAACCAACAGCTGGAAATGGACCCCTGTTCTCGACATATTTGTATGGA 2076  
QY 2974 ATCTTTTAAATACAGAAAAAGCCAAATCGGGTGCACAGTTTCAGTACTCATTTTCATGGATG 3033  
Db 2077 ATCTTTTAAATACAGAAAAAGCCAAATCGGGTGCACAGTTTCAGTACTCATTTTCATGGATG 2136  
QY 3034 AAGCTGACATTTCTGCAGATAGGAAGCTGTGATATACAGAAAGTCTGAAATGTGTTG 3093  
Db 2137 AAGCTGACATTTCTGCAGATAGGAAGCTGTGATATACAGAAAGTCTGAAATGTGTTG 2196  
QY 3094 GTTCTTCAATGTTTCTCAAAAGTAAATGGGGGATCGGCTACCGCCTGAGCATGTACATAG 3153

Db 2197 GTTCTTCAATGTTTCTCAAAAGTAAATGGGGATCGGCTACCGCCTGAGCATGTACATAG 2256  
QY 3154 ACAAATATTTTGTGCCACAGAAATCTCTTTCTTCTACTGGTTTAAACAAATATACCTGGAGCTA 3213  
Db 2257 ACAAATATTTTGTGCCACAGAAATCTCTTTCTTCTACTGGTTTAAACAAATATACCTGGAGCTA 2316  
QY 3214 CTTTATTAACAACAGAAATGACCAACAACTTTGTGTATAGCTTTGCGCTTTCAAGGACATGACGA 3273  
Db 2317 CTTTATTAACAACAGAAATGACCAACAACTTTGTGTATAGCTTTGCGCTTTCAAGGACATGACGA 2376  
QY 3274 AATTTTCAGGTTTGTCTTCTGCCCTTAGACAGTCAATTCAAATTTGGGTGTCAATTTCTTATG 3333  
Db 2377 AATTTTCAGGTTTGTCTTCTGCCCTTAGACAGTCAATTCAAATTTGGGTGTCAATTTCTTATG 2436  
QY 3334 GTGTTTCCATGACGACTTTTGGGAAGCGTATTTTAAAGCTAGAAAGTTGAAGCAGAAATTTG 3393  
Db 2437 GTGTTTCCATGACGACTTTTGGGAAGCGTATTTTAAAGCTAGAAAGTTGAAGCAGAAATTTG 2496  
QY 3394 ACCAAGCAGATATATAGTGTATTTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAAT 3453  
Db 2497 ACCAAGCAGATATATAGTGTATTTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAAT 2556  
QY 3454 CTTTGTATGAAATGGAAACAGAGCTTACTTATTTCTTTCTGAAACCAAGGCTTCTCTAGTGA 3513  
Db 2557 CTTTGTATGAAATGGAAACAGAGCTTACTTATTTCTTTCTGAAACCAAGGCTCTCTAGTGA 2616  
QY 3514 GCACCATGAGCCTTTTGGAAACCAACAGATGTATACAAATAGCAAAAGTTTCATTTCTTTACCT 3573  
Db 2617 GCACCATGAGCCTTTTGGAAACCAACAGATGTATACAAATAGCAAAAGTTTCATTTCTTTACCT 2676  
QY 3574 TGAACCGTGAAGTAAATCAGTGAGATCAGTGTGTTCTTCTGCTTTTAAATTTTTTTCACAG 3633  
Db 2677 TGAACCGTGAAGTAAATCAGTGAGATCAGTGTGTTCTTCTGCTTTTAAATTTTTTTCACAG 2736  
QY 3634 TTCAGATTTTATGTTTGTGTTTTCATCTCTTTTAAAAATGCTGTGGTTTCCCATCAAAAC 3693  
Db 2737 TTCAGATTTTATGTTTGTGTTTTCATCTCTTTTAAAAATGCTGTGGTTTCCCATCAAAAC 2796  
QY 3694 TTGTTTCCAGACTTATATTTTCTAAACCTGGAGACAAACCAATATAACAAACAAAGTC 3753  
Db 2797 TTGTTTCCAGACTTATATTTTCTAAACCTGGAGACAAACCAATATAACAAACAAAGTC 2856  
QY 3754 TGCTTCTTCAAAATTTCTGCT-----GACTCAGATATCAGTGATCTTA 3795  
Db 2857 TGCTTCTTCAAAATTTCTGCTGGTGAGAGTGTWNGTGAAGCTCAGATATACAGTGATCTTA 2916  
QY 3796 TTAGCTTTTTCACAGCCAGAAACATAATGGTGAACGATTAATGACAGTGAATGTAT 3855  
Db 2917 TTAGCTTTTTCACAGCCAGAAACATAATGGTGAACGATTAATGACAGTGAATGTAT 2976  
QY 3856 CCGTGGCTCCCATAGTGGGCTTTTAAATGTGATGCAATTCAGAAAGGACTATGTTTTG 3915  
Db 2977 CCGTGGCTCCCATAGTGGGCTTTTAAATGTGATGCAATTCAGAAAGGACTATGTTTTG 3036  
QY 3916 CAGCTGTTTTCAACAGTACTATGTTTATTTTACTATATATAGTGAATATCATTTAGTA 3975  
Db 3037 CAGCTGTTTTCAACAGTACTATGTTTATTTTACTATATATAGTGAATATCATTTAGTA 3096  
QY 3976 ACTTACTATCTTTATCAATTTTAAATGTGACTGAAACCAATCCAGATCTCGAGTACCCATCT 4035  
Db 3097 ACTTACTATCTTTATCAATTTTAAATGTGACTGAAACCAATCCAGATCTCGAGTACCCATCT 3156  
QY 4036 TTCAGAAATTTACTGATATAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTCCTTG 4095  
Db 3157 TTCAGAAATTTACTGATATAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTCCTTG 3216  
QY 4096 G-AATCATTTGTTACTGCAATGCCACTTACTTTTGGCATGAAATTCGAGAAATCATAG 4154  
Db 3217 GAAATCATTTGTTACTGCAATGCCACTTACTTTTGGCATGAAATTCGAGAAATCATAG -- 3274  
QY 4155 ATCAAAAGCTTATATCTCAACTTAACTTTTCAGGCTTTTTCGCCATCTGCATATTTGGATTGGA 4214

Db 3275 -----AGGCTTTTGGCACTGTCATATTGGATTGGA 3305  
Qy 4215 CAAGCTGTTGTGATATCCCTTATTTTATATCACTTATTTTGTGATGCTAGGAAGCTTA 4274  
Db 3306 CAAGCTGTTGTGATATCCCTTATTTTATCACTTATTTTGTGATGCTAGGAAGCTTA 3365  
Qy 4275 CTGGCAATTCATATGATATATTTTATATCTGTAAGTTCCTTGTGTGGTTTTTTCG 4334  
Db 3366 TTGGCAATTCATATGATATATTTTATATCTGTAAGTTCCTTGTGTGGTTTTTTCG 3425  
Qy 4335 CTTATTGGTATGTTTCATCAGTTATTTCTGTTCATCTATATATGTTCTTTTCACTTTAAG 4394  
Db 3426 CTTATTGGTATGTTTCATCAGTTATTTCTGTTCATCTATATATGTTCTTTTCACTTTAAG 3485  
Qy 4395 AAAATTTTAAATACCAAGAAATTTTGTGCTATTTATCTGTCGAGCGTTGNCCTGT 4454  
Db 3486 AAAATTTTAAATACCAAGAAATTTTGTGCTATTTATCTGTCGAGCGTTGNCCTGT 3545  
Qy 4455 ATTGCAATCACTGAATAAATCTTTCTTTATGGGATACACAATTCGAATTTCTTTCAATTAT 4514  
Db 3546 ATTGCAATCACTGAATAAATCTTTCTTTATGGGATACACAATTCGAATTTCTTTCAATTAT 3605  
Qy 4515 GCCTTTTGTATCATCATTCOAATCTATCCACTTCTAGGTTGCCGTGATTTCTTTCAATAAG 4574  
Db 3606 GCCTTTTGTATCATCATTCOAATCTATCCACTTCTAGGTTGCCGTGATTTCTTTCAATAAG 3665  
Qy 4575 ATTCTTTGGAGATGTACGAATAAATGTGACACCTATATATCCATGGGATAGCTTTCA 4634  
Db 3666 ATTCTTTGGAGAAATGTACGAATAAATGTGACACCTATATATCCATGGGATAGCTTTCA 3725  
Qy 4635 GTAGCTGTATATCGCCTTACCTGACGTGTACTGTGGATTTTCTCTTCAATAACTAT 4694  
Db 3726 GTAGCTGTATATCGCCTTACCTGACGTGTACTGTGGATTTTCTCTTCAATAACTAT 3785  
Qy 4695 GAGAAAATAATATGAGGAGATCAATAAAGAAAAGATCCCTTTTTCAGAAAACCTTTCAACG 4754  
Db 3786 GAGAAAATAATATGAGGAGATCAATAAAGAAAAGATCCCTTTTTCAGAAAACCTTTCAACG 3845  
Qy 4755 AAGTCTAAATAATAGGAGCTTCAGAACCCAGACCAATCAGGATGAAGATGATGTC 4814  
Db 3846 AAGTCTAAATAATAGGAGCTTCAGAACCCAGACCAATCAGGATGAAGATGAAGATGTC 3905  
Qy 4815 AAAGCTGAAAGACTTAAAGGTCGAAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAACCA 4874  
Db 3906 AAAGCTGAAAGACTTAAAGGTCGAAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAACCA 3965  
Qy 4875 TCCATTATGTCAGCAATTTGCAATAAAGATATGATGACAAAGAAAGATTTTCTTTTCA 4934  
Db 3966 TCCATTATGTCAGCAATTTGCAATAAAGATATGATGACAAAGAAAGATTTTCTTTTCA 4025  
Qy 4935 AGAAAGTAAAGAGTGGCAACTAAATACATCTTCTGTCGTAAGAAAGGAGATC 4994  
Db 4026 AGAAAGTAAAGAGTGGCAACTAAATACATCTTCTGTCGTAAGAAAGGAGATC 4085  
Qy 4995 TTAGGACTATTGGGTCGCAATTTGTCGCAAAAGCAAAATTTATATTTCTGTTGGT 5054  
Db 4086 TTAGGACTATTGGGTCGCAATTTGTCGCAAAAGCAAAATTTATATTTCTGTTGGT 4145  
Qy 5055 GATATTGAAACCACTTCAGGCCAGGATTTTATAGGAGATTTCTTTCAGAGACAAGTGAA 5114  
Db 4146 GATATTGAAACCACTTCAGGCCAGGATTTTATAGGAGATTTCTTTCAGAGACAAGTGAA 4205  
Qy 5115 GATGATGATTCACGTAAGTGTATGGGTTACTGTCTCAGATAAACCTTTTGGCCAGAT 5174  
Db 4206 GATGATGATTCACGTAAGTGTATGGGTTACTGTCTCAGATAAACCTTTTGGCCAGAT 4265  
Qy 5175 ACTACATTTGAGGAACATTTTGAATTTATGGAGCTGTCAAAGGAATGAGTGAAGTGAC 5234  
Db 4266 ACTACATTTGAGGAACATTTTGAATTTATGGAGCTGTCAAAGGAATGAGTGAAGTGAC 4325  
Qy 5235 ATGAAAGAGTCAATAAGTCGAATAACACATGCACTTGAATTTAAAGAACATCTTTCAGAG 5294  
Db 4326 ATGAAAGAGTCAATAAGTCGAATAACACATGCACTTGAATTTAAAGAACATCTTTCAGAG 4385

Qy 5295 ACTGTAAGAAACTACCTGCAGGAATCAAAAGAAAGTTGTGTTTGTCTCTAAGTATGCTA 5354  
Db 4386 ACTGTAAGAAACTACCTGCAGGAATCAAAAGAAAGTTGTGTTTGTCTCTAAGTATGCTA 4445  
Qy 5355 GGAATTCCTCAGATTACCTTTGTCTAGATGAACCATCTACAGGTATGGATCCCAAGCCAAA 5414  
Db 4446 GGAATTCCTCAGATTACCTTTGTCTAGATGAACCATCTACAGGTATGGATCCCAAGCCAAA 4505  
Qy 5415 CAGCAGATGTCGAGAGCAATTCGAACTGATTTAAAAACAGAAAGCGGCTCTATTCTG 5474  
Db 4506 CAGCAGATGTCGAGAGCAATTCGAACTGATTTAAAAACAGAAAGCGGCTCTATTCTG 4565  
Qy 5475 ACCACTCACTATATCGAGAGGAGCAGAGCTGTCTGTGATCGAGTAGCTATCATGTGTCT 5534  
Db 4566 ACCACTCACTATATCGAGAGGAGCAGAGCTGTCTGTGATCGAGTAGCTATCATGTGTCT 4625  
Qy 5535 GGGCAGTTAAGATGTATCGGAACAGTACAACATCTAAAGAGTAATTTTGGAAAAGGCTAC 5594  
Db 4626 GGGCAGTTAAGATGTATCGGAACAGTACAACATCTAAAGAGTAATTTTGGAAAAGGCTAC 4685  
Qy 5595 TTTTGGAAATTTAAATTTGAAGACTGGATAGAAAACCTAGAAAGTAGACCGCTTCAAGA 5654  
Db 4686 TTTTGGAAATTTAAATTTGAAGACTGGATAGAAAACCTAGAAAGTAGACCGCTTCAAGA 4745  
Qy 5655 GAAATTCAGTATATTTTCCAAATGCAAGCCGTCAAGAAAGTTTCTTCTATTTTGGCT 5714  
Db 4746 GAAATTCAGTATATTTTCCAAATGCAAGCCGTCAAGAAAGTTTCTTCTATTTTGGCT 4805  
Qy 5715 TATAAAATTCCTAAGAAAGATGTTCAAGTCTTTCACAACTTTTTTAAAGCTGGAAGAA 5774  
Db 4806 TATAAAATTCCTAAGAAAGATGTTCAAGTCTTTCACAACTTTTTTAAAGCTGGAAGAA 4865  
Qy 5775 GCTAAACATGCTTTTGGCAATTTGAAGATATAGCTTTTCTCAAGCAACATTTGAACAGGTT 5834  
Db 4866 GCTAAACATGCTTTTGGCAATTTGAAGATATAGCTTTTCTCAAGCAACATTTGAACAGGTT 4925  
Qy 5835 TTTGTAGAACTCCTTAAAGAAACAGAGGAGGAAGATAATAGTTGTGGAACCTTTAAACAGC 5894  
Db 4926 TTTGTAGAACTCCTTAAAGAAACAGAGGAGGAAGATAATAGTTGTGGAACCTTTAAACAGC 4985  
Qy 5895 ACATTTTGGTGGAAACGAAACAGAAAGATAGATGATTTTGAATTTGTATTTGTTCCG 5954  
Db 4986 ACATTTTGGTGGAAACGAAACAGAAAGATAGATGATTTTGAATTTGTATTTGTTCCG 5045  
Qy 5955 TCTGCTTACTGGGACTTCTTCTTCTTCACTTAACTTTTAACTTTGTTTAAAGAGTTT 6014  
Db 5046 TCTGCTTACTGGGACTTCTTCTTCTTCACTTAACTTTTAACTTTGTTTAAAGAGTTT 5105  
Qy 6015 TATTGGAATGGTAACTGAGAAACCAAGAACGCACTTGAAATTTTCTAAGCTCCTTAAAT 6074  
Db 5106 TATTGGAATGGTAACTGAGAAACCAAGAACGCACTTGAAATTTTCTAAGCTCCTTAAAT 5165  
Qy 6075 GAAATGCTGTGGTGTGTGTTTTGTTTTCTTTTAAATAAAAGCTATGATATAAATAGTGA 6134  
Db 5166 GAAATGCTGTGGTGTGTGTTTTGTTTTCTTTTAAATAAAAGCTATGATATAAATAGTGA 5225  
Qy 6135 AGCTGCATGTTTGTATTTGAGTATATGAACTATATAGTTTGTATGTATGTTTTCACC 6194  
Db 5226 AGCTGCATGTTTGTATTTGAGTATATGAACTATATAGTTTGTATGTATGTTTTCACC 5285  
Qy 6195 ATTCAGAAACAGTGTCTGTAATTTTGAATTTTAAAGGAATTTGTAATAGAAATAGTTTAT 6254  
Db 5286 ATTCAGAAACAGTGTCTGTAATTTTGAATTTTAAAGGAATTTGTAATAGAAATAGTTTAT 5345  
Qy 6255 TTTTAAAGTTATCTTTAAGTTTTATGCCATCTTCTTAAATAAGTAGCGTAATGTTCCAACTAA 6314  
Db 5346 TTTTAAAGTTATCTTTAAGTTTTATGCCATCTTCTTAAATAAGTAGCGTAATGTTCCAACTAA 5405  
Qy 6315 ATAAAAAACTAATACATAACTTAATGTCATAGAAAAGATACATAAAGCAATGTGAAAGTT 6372  
Db 5406 ATAAAAAACTAATTCATAACTAATGTCATAGAAAAGATACATAAAGCAATGTGAAAGTT 5463

```
RESULT 6
AX417828
LOCUS AX417828 5262 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 5 from Patent WO0231147.
ACCESSION AX417828
VERSION AX417828.1 GI:21522944
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Hu, Y. and Nepomnichy, B.
AUTHORS Human transporter proteins and polynucleotides encoding the same
TITLE Patent: WO 0231147-A 5 18-APR-2002;
JOURNAL LEXICON GENETICS INC (US)
FEATURES
source location/Qualifiers
1..5262
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 1615 a 888 c 1012 g 1743 t 4 others
ORIGIN
Query Match 80.4%; Score 5247.4; DB 6; Length 5262;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 5248; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY 757 ACTGTTGATATGGTGGTATTTCAAATTCGGTCTACCCCTATTTCAACAGCTGTTTACT 816
DB 1 ACTGTTGATATGGTGGTATTTCAAATTCGGTCTACCCCTATTTCAACAGCTGTTTACT 60
QY 817 TTTCAGAGCTGACAGATGCTGCCATGCTGCTCCAGTCTGCTCCAGTCTTCTTAAGAGAGACGCT 876
DB 61 TTTCAGAGCTGACAGATGCTGCCATGCTGCTCCAGTCTGCTCCAGTCTTCTTAAGAGAGACGCT 120
QY 877 TGGAGTATGCTTAATCCATCTTACCTGGGAGTGAACAGAGCTGCTTATTTGCGGTAAAAA 936
DB 121 TGGAGTATGCTTAATCCATCTTACCTGGGAGTGAACAGAGCTGCTTATTTGCGGTAAAAA 180
QY 937 ATTACATGATGTTTACGCGTGGCTCCGGTGTGTTGTTGTTGTTTCTTTTAAATGAG 996
DB 181 ATTACATGATGTTTACGCGTGGCTCCGGTGTGTTGTTGTTGTTTCTTTTAAATGAG 240
QY 997 TTATTATCAGAAACATGCTCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACCAGAA 1056
DB 241 TTATTATCAGAAACATGCTCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACCAGAA 300
QY 1057 CACTTCTACTGAAGAATTACTTAATTAATGACAGAACCAAAAGAGTAGTGTTCAGGAAA 1116
DB 301 CACTTCTACTGAAGAATTACTTAATTAATGACAGAACCAAAAGAGTAGTGTTCAGGAAA 360
QY 1117 TCTTTTTCCACTATTTTTTTTATTTGGTTAATTAATTAATTAATTAATTAATTAATTAATTA 1176
DB 361 TCTTTTTCCACTATTTTTTTTATTTGGTTAATTAATTAATTAATTAATTAATTAATTAATTA 420
QY 1177 AGAAATATGAAGAGTGCCTTAATATAGAACTCAATCTATGGAAGTGTCTTCTTCTTA 1236
DB 421 AGAAATATGAAGAGTGCCTTAATATAGAACTCAATCTATGGAAGTGTCTTCTTCTTA 480
QY 1237 ATCTAAATCTTGGATATACCTCCAGTGTACTAATTAATTAATTAATTAATTAATTAATTA 1296
DB 481 ATCTAAATCTTGGATATACCTCCAGTGTACTAATTAATTAATTAATTAATTAATTAATTA 540
QY 1297 CTACTGATCATCTACCTGATGTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1356
DB 541 CTACTGATCATCTACCTGATGTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
QY 1357 TAACATCCAGTCTCTTAAGCCGAGCACTTTGTAGTGTGTTTCAAGAGCTCCATGT 1416
DB 601 TAACATCCAGTCTCTTAAGCCGAGCACTTTGTAGTGTGTTTCAAGAGCTCCATGT 660
1417 CCTATGAACCTTCGTTTTTCTCTGATATGATTCAGGATATCTTCTATTTATATGATGATCAA 1476
661 CCTATGAACCTTCGTTTTTCTCTGATATGATTCAGGATATCTTCTATTTATATGATGATCAA 720
1477 GAGCTGGCTGTTTCAAAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGGTTTTCACAGTTT 1536
721 GAGCTGGCTGTTTCAAAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGGTTTTCACAGTTT 780
1537 TACAAGCATCATAGATGCTGCCATATATACAGTTTGAAGACCAATTTCTCTTTTGGAGG 1596
781 TACAAGCATCATAGATGCTGCCATATATACAGTTTGAAGACCAATTTCTCTTTTGGAGG 840
1597 AGCTGGAGTCACTAAAGCTGTTTATATGGAGAACTGCTGTTGTAGAAATAGATACCT 1656
841 AGCTGGAGTCACTAAAGCTGTTTATATGGAGAACTGCTGTTGTAGAAATAGATACCT 900
1657 TTCCCGGAGGAGTAATTTTAATATACCTAGTTATAGCAATTTTCCCTTTTGGATACCTTT 1716
901 TTCCCGGAGGAGTAATTTTAATATACCTAGTTATAGCAATTTTCCCTTTTGGATACCTTT 960
1717 TGGCAATTCATATCGTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGATAATGG 1776
961 TGGCAATTCATATCGTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGATAATGG 1020
1777 GACTTTCATGATGCTGCTTTTGGCTTTTCTCTGCTTCTCTATATACAAAGTTTAAATTTTC 1836
1021 GACTTTCATGATGCTGCTTTTGGCTTTTCTCTGCTTCTCTATATACAAAGTTTAAATTTTC 1080
1837 TTATGCTCCTTTTATGGAGTCAATTTGGGACAGCTTCTTTGTTTATTTTCTCAAAGTAGCA 1896
1081 TTATGCTCCTTTTATGGAGTCAATTTGGGACAGCTTCTTTGTTTATTTTCTCAAAGTAGCA 1140
1897 GCATTCGTATATTTCTGCTTTTCTTTATGCTGATATCATCTGATTTTCTTTTCTTTAA 1956
1141 GCATTCGTATATTTCTGCTTTTCTTTATGCTGATATCATCTGATTTTCTTTTCTTTAA 1200
1957 TGCTGACACCTCTTTTAAATAATCAAAACATGCTGGGAAATAGTTGAATTTTGTGTACTG 2016
1201 TGCTGACACCTCTTTTAAATAATCAAAACATGCTGGGAAATAGTTGAATTTTGTGTACTG 1260
2017 TGCTTTTGGATTTATTTGGCTTTATGATTAATCTCTATAGAAAGTTTTCCTCAAATCGTTAG 2076
1261 TGCTTTTGGATTTATTTGGCTTTATGATTAATCTCTATAGAAAGTTTTCCTCAAATCGTTAG 1320
2077 TGCTGCTTTTCTGCTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2136
1321 TGCTGCTTTTCTGCTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
2137 ATTTAGAAGATTTTAAATGAAGTGTCTTCAATTTTCAAAATTTGACTGAGGCGCCATATCCTC 2196
1381 ATTTAGAAGATTTTAAATGAAGTGTCTTCAATTTTCAAAATTTGACTGAGGCGCCATATCCTC 1440
2197 TAATTAATCAATTTATCATGCTCACACTTAATAGTATATTTCTATGCTCTCTTGGCTGCTCT 2256
1441 TAATTAATCAATTTATCATGCTCACACTTAATAGTATATTTCTATGCTCTCTTGGCTGCTCT 1500
2257 ATCTTCATCAAGTCAATTTCCAGGGAAATTTGGCTTACGGAGATCATCTTTATATTTCTGA 2316
1501 ATCTTCATCAAGTCAATTTCCAGGGAAATTTGGCTTACGGAGATCATCTTTATATTTCTGA 1560
2317 AGCTTTCAATTTGGTCAAAAGAGTAAAGAAATTTATGAGGAGTTTATCAGAGGCAATGTTA 2376
1561 AGCTTTCAATTTGGTCAAAAGAGTAAAGAAATTTATGAGGAGTTTATCAGAGGCAATGTTA 1620
2377 ATGGAATATTTAGTTTGTAGTAAATTTATGAGCAGTTTCTTCCAGAAATTTGTAGGAAAG 2436
1621 ATGGAATATTTAGTTTGTAGTAAATTTATGAGCAGTTTCTTCCAGAAATTTGTAGGAAAG 1680
2437 AAGCCATAGATTTAGTGTATTTCAAGACATACAGAAAGAGGTTGAAATGTGGAGG 2496
1681 AAGCCATAGATTTAGTGTATTTCAAGACATACAGAAAGAGGTTGAAATGTGGAGG 1740
2497 CTTTGAAGAAATTTGTCAATTTGACATATATGAGGAGTTCAGATTACTGCTTACTTTGGCCACA 2556
```

|||||  
1741 CTTTGAGAAATTTGTCTATTGTGACATATATGAGGGTCAGATTACTGCTTACTTGGCCACA 1800  
|||||  
2557 GTGGAAACAGAAAGAGTACATTGATGAATATCTTTGTGGACCTCTGCCACACCTTCTGTATG 2616  
|||||  
1801 GTGGAAACAGAAAGAGTACATTGATGAATATCTTTGTGGACCTCTGCCACACCTTCTGTATG 1860  
|||||  
2617 GGTTCGATCTATATATGGACAGAGCTCTCAGAAATAGATGAATGTTTGAAGCAAGAA 2676  
|||||  
1861 GGTTCGATCTATATATGGACAGAGCTCTCAGAAATAGATGAATGTTTGAAGCAAGAA 1920  
|||||  
2677 AAATGATTTGGCAATTTGTCCACAGTTAGATATACACTTTGATGTTTTCAGAGTAGAAGAAA 2736  
|||||  
1921 AAATGATTTGGCAATTTGTCCACAGTTAGATATACACTTTGATGTTTTCAGAGTAGAAGAAA 1980  
|||||  
2737 ATTTATCAATTTTGGCTTCAATCAAGGGATACACAGCAACAATATAAATACAAGAGTGC 2796  
|||||  
1981 ATTTATCAATTTTGGCTTCAATCAAGGGATACACAGCAACAATATAAATACAAGAGTGC 2040  
|||||  
2797 AGAAGGTTTTACTAGATTTAGACATGACAGACTATCAAGATACCAAGCTTAAAAAATTAA 2856  
|||||  
2041 AGAAGGTTTTACTAGATTTAGACATGACAGACTATCAAGATACCAAGCTTAAAAAATTAA 2100  
|||||  
2857 GTGGTGTCAAAAAGAAAGCTGTCTATTAGGAATTCGTCTTTGGGAACCCCAAGATAC 2916  
|||||  
2101 GTGGTGTCAAAAAGAAAGCTGTCTATTAGGAATTCGTCTTTGGGAACCCCAAGATAC 2160  
|||||  
2917 TGCTGTAGATGAACCAAGCTGGAAATGGACCCCTGTCTCGACATATGTTATGAATC 2976  
|||||  
2161 TGCTGTAGATGAACCAAGCTGGAAATGGACCCCTGTCTCGACATATGTTATGAATC 2220  
|||||  
2977 TTTTAAATACAGAAAGCAATCGGTGACAGTGTTCAGTACTCATTTTCATGATGAAG 3036  
|||||  
2221 TTTTAAATACAGAAAGCAATCGGTGACAGTGTTCAGTACTCATTTTCATGATGAAG 2280  
|||||  
3037 CTGACATTTCTGACATAGAAAGCTGTGATATACAAAGGAATGCTGAAATGTTGGTT 3096  
|||||  
2281 CTGACATTTCTGACATAGAAAGCTGTGATATACAAAGGAATGCTGAAATGTTGGTT 2340  
|||||  
3097 CTTCAATGTTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGGATGTACATAGACA 3156  
|||||  
2341 CTTCAATGTTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACA 2400  
|||||  
3157 AATATTGTCACAGATCTCTTTCTTCACTGTTTAAACAAATATACCTGGAGCTACTT 3216  
|||||  
2401 AATATTGTCACAGATCTCTTTCTTCACTGTTTAAACAAATATACCTGGAGCTACTT 2460  
|||||  
3217 TATTACAAAGAAATGACCAACAACTTGTGTATAGCTTTGCCCTTTCAAGGACATGGACAAAT 3276  
|||||  
2461 TATTACAAAGAAATGACCAACAACTTGTGTATAGCTTTGCCCTTTCAAGGACATGGACAAAT 2520  
|||||  
3277 TTTTCAGGTTGTTTCTGCGCTAGACAGTCAATTCAAATTTGGGTGTCAATTTCTTATGGTG 3336  
|||||  
2521 TTTTCAGGTTGTTTCTGCGCTAGACAGTCAATTCAAATTTGGGTGTCAATTTCTTATGGGG 2580  
|||||  
3337 TTTTCAGTACAGCTTTTGGAGAGCTATTTTAAAGCTAGAGTTTGAAGCAGAAATTTGACC 3396  
|||||  
2581 TTTTCAGTACAGCTTTTGGAGAGCTATTTTAAAGCTAGAGTTTGAAGCAGAAATTTGACC 2640  
|||||  
3397 AAGCAGATTTATGTTATTTACTCAGCAGCCACTGGAGGAGAAATGGATTCAAAATCTTT 3456  
|||||  
2641 AAGCAGATTTATGTTATTTACTCAGCAGCCACTGGAGGAGAAATGGATTCAAAATCTTT 2700  
|||||  
3457 TTGATGAATTTGGAAACAGAGCTTACTTCTTCTGAAACCAAGCTTCTCTAGTGAGCA 3516  
|||||  
2701 TTGATGAATTTGGAAACAGAGCTTACTTCTTCTGAAACCAAGCTTCTCTAGTGAGCA 2760  
|||||  
3517 CCATGAGCCTTTGGAAACCAAGATGTATCAATAGCAAAAGTTTCATTTCTTTACCTTGA 3576  
|||||  
2761 CCATGAGCCTTTGGAAACCAAGATGTATCAATAGCAAAAGTTTCATTTCTTTACCTTGA 2820  
|||||  
3577 AACGTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTC 3636  
|||||

2821 AACGTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTC 2880  
|||||  
3637 AGATTTTATGTTTGTGTTTGTTCATCACTCTTTTAAAAATGCTGTGTTCCCATCAAACTTG 3696  
|||||  
2881 AGATTTTATGTTTGTGTTTGTTCATCACTCTTTTAAAAATGCTGTGTTCCCATCAAACTTG 2940  
|||||  
3697 TTTCCAGACTTATATTTTCTAAAACTTGAGACAAACCAACATATAAATACAAAACTGCTGC 3756  
|||||  
2941 TTTCCAGACTTATATTTTCTAAAACTTGAGACAAACCAACATATAAATACAAAACTGCTGC 3000  
|||||  
3757 TTTCTCAAAATTTCTGCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAGCCAGA 3816  
|||||  
3001 TTTCTCAAAATTTCTGCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAGCCAGA 3060  
|||||  
3817 ACATAATGCTGACGATGATTAATGACAGTGATATGTTATCCGTGCTCTCCCATAGTGGG 3876  
|||||  
3061 ACATAATGCTGACGATGATTAATGACAGTGATATGTTATCCGTGCTCTCCCATAGTGGG 3120  
|||||  
3877 CTTTAAATGTGATGCAATTCAGAAAAAGGACTATGTTTTCAGAGCTGTTTTCACAGTACTA 3936  
|||||  
3121 CTTTAAATGTGATGCAATTCAGAAAAAGGACTATGTTTTCAGAGCTGTTTTCACAGTACTA 3180  
|||||  
3937 TGGTTTATCTTTACTTATATTTAGTGAATATCAATTAGTAACTACTATCTTTTATCAATTAA 3996  
|||||  
3181 TGGTTTATCTTTACTTATATTTAGTGAATATCAATTAGTAACTACTATCTTTTATCAATTAA 3240  
|||||  
3997 ATGTGACTGAAACCAATCCAGATCTGAGTACCCCACTCTTTCAAGAAATTAATCTGATATAG 4056  
|||||  
3241 ATGTGACTGAAACCAATCCAGATCTGAGTACCCCACTCTTTCAAGAAATTAATCTGATATAG 3300  
|||||  
4057 TTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCAATGTTATGCTCAATGC 4116  
|||||  
3301 TTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCAATGTTATGCTCAATGC 3360  
|||||  
4117 CACCTTACTTTGCCATGAAATGCAGAGATCATAGATCAAGCTTATATCTGAACTTA 4176  
|||||  
3361 CACCTTACTTTGCCATGAAATGCAGAGATCATAGATCAAGCTTATATCTGAACTTA 3420  
|||||  
4177 AACTTTTCAAGCTTTTGGCATCTGCATATGGATGGACAAGCTGTTGTTGATATCCCT 4236  
|||||  
3421 AACTTTTCAAGCTTTTGGCATCTGCATATGGATGGACAAGCTGTTGTTGATATCCCT 3480  
|||||  
4237 TATTTTATCATCTTATTTTGTATGCTGAGAGCTTACTGGCATTTTCAATTAAGATTAT 4296  
|||||  
3481 TATTTTATCATCTTATTTTGTATGCTGAGAGCTTATGGCATTTTCAATTAAGATTAT 3540  
|||||  
4297 ATTTTATACGTAAAGTTCCTTGTGCTGGTTTTTGGCTTATGCTTATGTTTCCATCAG 4356  
|||||  
3541 ATTTTATACGTAAAGTTCCTTGTGCTGGTTTTTGGCTTATGCTTATGTTTCCATCAG 3600  
|||||  
4357 TTTATCTGTTCACTTATATTTGCTTCTTTTCAAGAAATTTTAAATACCAAGAAAT 4416  
|||||  
3601 TTTATCTGTTCACTTATATTTGCTTCTTTTCAAGAAATTTTAAATACCAAGAAAT 3660  
|||||  
4417 TTTGTTCACTTATCTTATCTGTCGAGCTTGNCTTGTATGTCATCTCACTGAAATAACTTT 4476  
|||||  
3661 TTTGTTCACTTATCTTATCTGTCGAGCTTGGCTTGTATGTCATCTCACTGAAATAACTTT 3720  
|||||  
4477 TCTTTATGGGATACAAATTCGAATCTTCTCATATATGCTTTTGTATCATCATTTCCAA 4536  
|||||  
3721 TCTTTATGGGATACAAATTCGAATCTTCTCATATATGCTTTTGTATCATCATTTCCAA 3780  
|||||  
4537 TCTATCCACTTCTAGCTTGGCTGATTTCTTTTCAAAAGATTTCTTGGAAAGAAATGACGAA 4596  
|||||  
3781 TCTATCCACTTCTAGCTTGGCTGATTTCTTTTCAAAAGATTTCTTGGAAAGAAATGACGAA 3840  
|||||  
4597 AAAATGTGACACCTTATAATCCATGGGATAGCTTTTCACTAGTCTGTTATATCGCTTACC 4656  
|||||  
3841 AAAATGTGACACCTTATAATCCATGGGATAGCTTTTCACTAGTCTGTTATATCGCTTACC 3900  
|||||  
4657 TGCAGTGTGATCTGTGGATTTTCTCTTACAATATCTATGAGAAAAAATATGAGGCGAGAT 4716  
|||||  
3901 TGCAGTGTGATCTGTGGATTTTCTCTTACAATATCTATGAGAAAAAATATGAGGCGAGAT 3960  
|||||

```
QY 4717 CAATAAGAAAGATCCCTTTTTCAGAAACCTTTCAACGAAGTCTAAAAATAGGAAGCTTC 4776
Db 3961 CAATAAGAAAGATCCCTTTTTCAGAAACCTTTCAACGAAGTCTAAAAATAGGAAGCTTC 4020
QY 4777 CAGAACCCACAGACATAGAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGA 4836
Db 4021 CAGAACCCACAGACATAGAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGA 4080
QY 4837 AAGAGCTGATGGGTGGCAGTGTGGTGGAGAGAAACCATCATATATGGTGCAGCAATTTGC 4896
Db 4081 AAGAGCTGATGGGTGGCAGTGTGGTGGAGAGAAACCATCATATATGGTGCAGCAATTTGC 4140
QY 4897 ATAAAGATATGATGACAAGAAAGATTTCTTTTCAAGAAAGTAAAGAAAGTGGCAA 4956
Db 4141 ATAAAGATATGATGACAAGAAAGATTTCTTTTCAAGAAAGTAAAGAAAGTGGCAA 4200
QY 4957 CTAATAACATCTCTTTCTGTGTGAAAGAGAGATCTTAGACATATTTGGGTCCAAATG 5016
Db 4201 CTAATAACATCTCTTTCTGTGTGAAAGAGAGATCTTAGACATATTTGGGTCCAAATG 4260
QY 5017 GTCTGGCAAAAGCACAAATTAATTAATTTCTGTGTGGTGTATTTGAACCAACTTCAGGCC 5076
Db 4261 GTCTGGCAAAAGCACAAATTAATTAATTTCTGTGTGGTGTATTTGAACCAACTTCAGGCC 4320
QY 5077 AGGTATTTTTPAGGAGATTTATTTCTTCAGAGACAAGTGAAGATGATTTCACTGAAGTGA 5136
Db 4321 AGGTATTTTTPAGGAGATTTATTTCTTCAGAGACAAGTGAAGATGATTTCACTGAAGTGA 4380
QY 5137 TGGGTACTCTCTCAGATAAACCCTTTTGGGCCAGATACATTCAGAGAAACATTTTG 5196
Db 4381 TGGGTACTCTCTCAGATAAACCCTTTTGGGCCAGATACATTCAGAGAAACATTTTG 4440
QY 5197 AAATTTATGAGCTGCTCAAGAGATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 5256
Db 4441 AAATTTATGAGCTGCTCAAGAGATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 4500
QY 5257 TAACACATGCACTTGATTTTAAAGAAACATCTTCAGAGAGCTGTAAAGAAACTACCTGCAG 5316
Db 4501 TAACACATGCACTTGATTTTAAAGAAACATCTTCAGAGAGCTGTAAAGAAACTACCTGCAG 4560
QY 5317 GAATCAACGAAGTGTGGTGTCTTAAGTATGCTAGGAAATCCTCAGATTAATTTGC 5376
Db 4561 GAATCAACGAAGTGTGGTGTCTTAAGTATGCTAGGAAATCCTCAGATTAATTTGC 4620
QY 5377 TAGATGAACCATCTACAGGTATGATCCCAAGCCAAACAGCACATGTCGCGAGCAATTC 5436
Db 4621 TAGATGAACCATCTACAGGTATGATCCCAAGCCAAACAGCACATGTCGCGAGCAATTC 4680
QY 5437 GAATGCACTTAAAGAAAGCGGCTGCTATTCTGACCACTCACTATATGAGAGAGG 5496
Db 4681 GAATGCACTTAAAGAAAGCGGCTGCTATTCTGACCACTCACTATATGAGAGAGG 4740
QY 5497 CAGAGCTGTCTGTGATCGAGTATCATGTGTCTGGCAGTGAAGATGATCGGAA 5556
Db 4741 CAGAGCTGTCTGTGATCGAGTATCATGTGTCTGGCAGTGAAGATGATCGGAA 4800
QY 5557 CAGTACAACATCTAAAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTAAGG 5616
Db 4801 CAGTACAACATCTAAAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTAAGG 4860
QY 5617 ACTGGATAGAAAACTTAGAGTAGACCGCTTCAAGAGAGAAATTCAGTATATTTTCCCAA 5676
Db 4861 ACTGGATAGAAAACTTAGAGTAGACCGCTTCAAGAGAGAAATTCAGTATATTTTCCCAA 4920
QY 5677 ATGCAAGCCGTCAGGAAGTTTTCTTCTCTATTTTGGCTTATAAAATTCCTAAGGAAGTG 5736
Db 4921 ATGCAAGCCGTCAGGAAGTTTTCTTCTCTATTTTGGCTTATAAAATTCCTAAGGAAGTG 4980
QY 5737 TTCAAGTCCCTTTCACATCTTTTAAAGTGGAGAGCAATCAAGATGCTTTTGGCATTG 5796
Db 4981 TTCAAGTCCCTTTCACATCTTTTAAAGTGGAGAGCAATCAAGATGCTTTTGGCATTG 5040
```

```
QY 5797 AAGAATATAGCTTTTCTCAAGCAACATTCGAAACAGGTTTTTTGTAGAACTCACTAAAGAAC 5856
Db 5041 AAGAATATAGCTTTTCTCAAGCAACATTCGAAACAGGTTTTTTGTAGAACTCACTAAAGAAC 5100
QY 5857 AAGAGGAGAGCAATATAGTGTGGAACTTTTAAACAGCACATTTTGGTGGGACGAAAC 5916
Db 5101 AAGAGGAGAGCAATATAGTGTGGAACTTTTAAACAGCACATTTTGGTGGGACGAAAC 5160
QY 5917 AAGAAGATAGAGTAGTATTTTGAATTTGATTTGTTGCGTCTGCTTACTGGGACTTCTTTC 5976
Db 5161 AAGAAGATAGAGTAGTATTTTGAATTTGATTTGTTGCGTCTGCTTACTGGGACTTCTTTC 5220
QY 5977 TTTTTCACCTTAATTTAACTTTGGTTTAAAAAGTTTTT 6014
Db 5221 TTTTTCACCTTAATTTAACTTTGGTTTAAAAAGTTTTT 5258

RESULT 7
HSA275973
LOCUS
DEFINITION
Homo sapiens mRNA for ATP-binding cassette protein of the (ABCA
subfamily).
ACCESSION
AJ275973
VERSION
AJ275973.3 GI:22080663
KEYWORDS
ABCA subfamily; ATP-binding cassette protein.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Petry, F., Kotthaus, A. and Hirsch-Ernst, K.I.
Molecular cloning and tissue distribution of a novel ATP-binding
cassette (ABC) transporter belonging to the subfamily ABCA
Unpublished
2
Hirsch-Ernst, K.I.
Direct Submission
Submitted (02-MAR-2000) Hirsch-Ernst K.I., Department of
Toxicology, University of Goettingen, Robert-Koch-Strasse 40,
Lowery Saxony D-37075, GERMANY
Revised by [3]
3
(bases 1 to 5096)
Hirsch-Ernst, K.I.
Direct Submission
Submitted (20-NOV-2001) Hirsch-Ernst K.I., Department of
Toxicology, University of Goettingen, Robert-Koch-Strasse 40,
Lowery Saxony, D-37075, GERMANY
On Aug 1, 2002 this sequence version replaced gi:17046099.
LOCATION/Qualifiers
1..5096
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="testis"
66..4994
/function="ABC transporter"
/note="belongs to the ABCA subfamily"
/codon_start=1
/product="ATP-binding cassette protein"
/protein_id="CAB93535.3"
/db_xref="GI:22080664"

CDS
1..5096
/translation="MSTAIRVGVWROTFTLLKNVLIKCRTKKSVQVILPLPLFLF
WLLISMHPNKKYEEVFNIELNPMDFKTLNLLILGYPTVNTSSIMOKVSTDHLPD
VIITEYNNEKEMLTSSLSKSPNFVGVKPSVMSYELRFFDPMIPVSSIMYDSRAGCS
KSCAAQYNSGFTVLOASIDAAIIQLKTNVSLWKELESTKAVIMGETAVVIDTPR
GVLLYLVAFSPFGVFLAIIHVAEKKIKFELKIMGLHDTAFWLSWLLYLSLFL
NSLLMAVATASLLPQSSIVIFLLFLYGLSSVFFPALMLPLPKSKHVGIVPEFV
TVAFGEIGMILLIESFPKSLVWLPSPFCHCTFVIGIAQVHLEDFNEGASPNLTAG
PYPLITITMLTNSIFVYVLLAVILDQVIPGFGRRSLYFLKPSYWSKSRNBYEL
SEGVNNGNISFSEIIEPVSFEVGEAIRISGIQTKYRKKGNVEALRNLSFDIYEQ
ITALLGHSCTGLTLMNLICGLCPDPSDFASIVGHRVSEIDEMFEAKRMIGICPQIDI
HPDLVTVEENLSILASIKGIPANNIIQEVQKVLDDLDMDQTIKDNQAKLGGQKRKLS
```

LGIAVLGNPKILLIDBPTAGMDPCSRHIIWMNLKRYKANRVTVFSTHFDEADIIADR  
KAVISGMLKCGSSMFLKSGWIGYRLSNYIDKVCATBSLSLVKQHI PGATLIQDQ  
DOOLVYSLPKQMDKESGLFSLDLSHNSLGVISYGMTTLEDDVFLKLEVEAEIQQAD  
YSVFTQOPELEENDSKSPBEMEQSLIILSETKASLVSTMSLWKQMYIAKHPFTLK  
RESKRSVSVLLILLIIFFTQVIFMLFVHHSFKNAVPIKLVPDLYFLKPGDKPKHYKTS  
LLQNSADSISLISFPTSQNTIMVIMINDSDYVSAHSAALNMHSEKDYFAAFV  
NSTWVSLPILNIIISNYILHNLVETIQTIPFPBITDIFVKIELYFOAALIGI  
IVTAMPYPFAMENAEHKKIAYTQLKSLGLPSAYWIGQAVVDIPLFFIILMLGSL  
LAHYGLYFYTVPKFLAVCLIGYVPSVILFYIASPTFKKILNLTKEPWSFYLSVAAL  
ACIAITEIFPMGTITLIIHAYFCIIPIYPLGLCLISFIKLSKWNKVRKNDVTNPW  
DLASVAIISYLOQVIMIILLOYEKRYGSRIRKDPFFRNLSLTKSRKRLPEPPDNE  
DEDEDKAEKRLKELMGQCCKEKESIMVSNLHKEYDDKDFLLSRKRVKATYKIS  
FCVKGEIILLGPNAGKSTIINILVGIPEPTSGQIFLGDYSESSEDDSLKQWGY  
CPQINPLPDTLLOHEFIYGVKMGMSADMKVEI SRI THALDLKHEHLOKTVKLIYH  
IKKLCFLSMLGNPOTILLDEPSTGMDPKAKQHMRAIRTAFAKNEKRAAILTPAG  
EBAVCDRVAIMVSGQRCITGVOLHLSKFKGFIYFLEIKLWIEINLEVDRLQREIQY  
IFPNARQSFSSILAYIKPKEDVQSLSQSFPLEAKHAFALBEYSFSQATLEQVVF  
ELTKQEEDNSCGTLNLSLWERTQEDRVF"

BASE COUNT 1584 a 858 c 986 g 1668 t

ORIGIN

Query Match 77.2%; Score 5039; DB 9; Length 5096;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 5042; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 994 AGGTTTATTCAGAAACATCTGCACCTGCAATTAGGGAGGTAGGAGTTTGGAGACAGACCA 1053  
DB 49 AGGTTTATTCAGAAACATCTGCACCTGCAATTAGGGAGGTAGGAGTTTGGAGACAGACCA 108

QY 1054 GAACACTTCTACTCAAGAAATTACTTAATTAATGCAGAACCAAAAAGAGTAGTGTTCAGG 1113  
DB 109 GAACACTTCTACTCAAGAAATTACTTAATTAATGCAGAACCAAAAAGAGTAGTGTTCAGG 168

QY 1114 AAATCTTTTCCACTATTTTATTTGGTGAATATTAATTAATAGCATGCATCCAA 1173  
DB 169 AAATCTTTTCCACTATTTTATTTGGTGAATATTAATTAATAGCATGCATCCAA 228

QY 1174 ATAAGAAATAGAGAGTGCCTTAATAGAACTCAATCTATGGCAAGTTTACTCTTT 1233  
DB 229 ATAAGAAATAGAGAGTGCCTTAATAGAACTCAATCTATGGCAAGTTTACTCTTT 288

QY 1234 CTAATCTAATTTCTGGATATCTCCAGTGACTAATATTAACAAGCAGCATCATGCAGAAAG 1293  
DB 289 CTAATCTAATTTCTGGATATCTCCAGTGACTAATATTAACAAGCAGCATCATGCAGAAAG 348

QY 1294 TGTCTACTGATCATCTACCTGATGTCTAATATTAATGAGAAATATACAATGAAAGAA 1353  
DB 349 TGTCTACTGATCATCTACCTGATGTCTAATATTAATGAGAAATATACAATGAAAGAA 408

QY 1354 TGTTAACATCCAGTCTCTTAAGCCGAGCACTTGTAGGTGTGGTTTTCAGAGATCCCA 1413  
DB 409 TGTAAACATCCAGTCTCTTAAGCCGAGCACTTGTAGGTGTGGTTTTCAGAGATCCCA 468

QY 1414 TGTCTATGAATCTCGTTTCTTCTGATATGATTCAGATATCTTCTATTTATATGAAAT 1473  
DB 469 TGTCTATGAATCTCGTTTCTTCTGATATGATTCAGATATCTTCTATTTATGAAAT 528

QY 1474 CAAGAGCTGGCTGTTCAAAATCATGTGAGCTGCTCAGTACTGGTCTCAGGTTTTCACAG 1533  
DB 529 CAAGAGCTGGCTGTTCAAAATCATGTGAGCTGCTCAGTACTGGTCTCAGGTTTTCACAG 588

QY 1534 TTTTACAAGCATCATAGATGCTGCCAATATACAGTTGAGACCAATGTTTCTCTTGG 1593  
DB 589 TTTTACAAGCATCATAGATGCTGCCAATATACAGTTGAGACCAATGTTTCTCTTGG 648

QY 1594 AGGAGCTGGAGTCAACTAAAGCTGTTTATATGGAGAAACTGCTGTTGTGAGAAATAGATA 1653  
DB 649 AGGAGCTGGAGTCAACTAAAGCTGTTTATATGGAGAAACTGCTGTTGTGAGAAATAGATA 708

QY 1654 CTTTTCCTCCGAGGAGTAAATTTTAAATATACCTAGTATAGCAATTTTCACTTTTGGTACT 1713  
DB 709 CTTTTCCTCCGAGGAGTAAATTTTAAATATACCTAGTATAGCAATTTTCACTTTTGGTACT 768

QY 1714 TTTTGGCAATTCATATCCTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAA 1773  
DB 769 TTTTGGCAATTCATATCCTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAA 828

QY 1774 TGGGACTTCATGATACCTGCTTTTGGCTTTCTCTGGGTCTCTCTATATACAGTGTAAATTT 1833  
DB 829 TGGGACTTCATGATACCTGCTTTTGGCTTTCTCTGGGTCTCTCTATATACAGTGTAAATTT 888

QY 1834 TTTCTATGTCCTCTTATGCGCAGTCATTTGCGACAGCTTCTTTGTTATTTCTCTCAAGTA 1893  
DB 889 TTTCTATGTCCTCTTATGCGCAGTCATTTGCGACAGCTTCTTTGTTATTTCTCTCAAGTA 948

QY 1894 GCAGCAATTTGATATTTCTGCTTTTTTCTTTATGGAATATCATCTGTATTTTGTCTT 1953  
DB 949 GCAGCAATTTGATATTTCTGCTTTTTTCTTTATGGAATATCATCTGTATTTTGTCTT 1008

QY 1954 TAACTGCTGACACCTCTTTTAAAAAATCAAAACATGTCGGATAGTTGAAATTTTGTGTA 2013  
DB 1009 TAACTGCTGACACCTCTTTTAAAAAATCAAAACATGTCGGATAGTTGAAATTTTGTGTA 1068

QY 2014 CTGTGCTTTTGGATTTATTTGSCCTTATGATAATCTCTATAGAAAATTTTCCAAATCGT 2073  
DB 1069 CTGTGCTTTTGGATTTATTTGSCCTTATGATAATCTCTATAGAAAATTTTCCAAATCGT 1128

QY 2074 TAGTGTGGCTTTTCACTGCTTTTCTGCTCACTGTACTTTTGTGATGCTATTTGCACAGGTCA 2133  
DB 1129 TAGTGTGGCTTTTCACTGCTTTTCTGCTCACTGTACTTTTGTGATGCTATTTGCACAGGTCA 1188

QY 2134 TGCATTTAGAGATTTTAAAGAGTGTCTTCATTTTCAAAATTTGACTGCAGGCCCATATC 2193  
DB 1189 TGCATTTAGAGATTTTAAAGAGTGTCTTCATTTTCAAAATTTGACTGCAGGCCCATATC 1248

QY 2194 CTCTAATTTATACAAATATCATGCTCACACTTAAATAGTATATTTCTATGCTCTTTGGCTG 2253  
DB 1249 CTCTAATTTATACAAATATCATGCTCACACTTAAATAGTATATTTCTATGCTCTTTGGCTG 1308

QY 2254 TCTATCTTTGATCAAGTCAATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 2313  
DB 1309 TCTATCTTTGATCAAGTCAATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 1368

QY 2314 TGAAGCTTCATATTTGGTCAAGAGTAAAGAAATATGAGGAGTTATCAGAGGGCAATG 2373  
DB 1369 TGAAGCTTCATATTTGGTCAAGAGTAAAGAAATATGAGGAGTTATCAGAGGGCAATG 1428

QY 2374 TTAATGGAAATATTTAGTTTGTAGTAAATTTATTTGAGCAGTTTCTTTCAGAAATTTGTAGGAA 2433  
DB 1429 TTAATGGAAATATTTAGTTTGTAGTAAATTTATTTGAGCAGTTTCTTTCAGAAATTTGTAGGAA 1488

QY 2434 AAGAAAGCCATAAGAAATTTAGTGTATTTCAAGACACATACAGAAAGAGGTGAAAAATGTGG 2493  
DB 1489 AAGAAAGCCATAAGAAATTTAGTGTATTTCAAGACACATACAGAAAGAGGTGAAAAATGTGG 1548

QY 2494 AGGCTTTGAGAAATTTGTCAATTTGACATATATGAGGTCAGATTACTGCTTACTTGGCC 2553  
DB 1549 AGGCTTTGAGAAATTTGTCAATTTGACATATATGAGGTCAGATTACTGCTTACTTGGCC 1608

QY 2554 ACAGTGGACAGGAAGAGTACATTTGATGAATTTCTTTGTGAGCTCTGCCACCTTCTG 2613  
DB 1609 ACAGTGGACAGGAAGAGTACATTTGATGAATTTCTTTGTGAGCTCTGCCACCTTCTG 1668

QY 2614 ATGGGTTTGCATCTATATATGACACAGAGTCTCAGAAATAGATGAAATTTTGTGAGCAA 2673  
DB 1669 ATGGGTTTGCATCTATATATGACACAGAGTCTCAGAAATAGATGAAATTTTGTGAGCAA 1728

QY 2674 GAAAAATGATTTGGCAATTTGTGCCACAGTTAGATATACATTTGATGTTTGTGACAGTAGAG 2733  
DB 1729 GAAAAATGATTTGGCAATTTGTGCCACAGTTAGATATACATTTGATGTTTGTGACAGTAGAG 1788

QY 2734 AAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAACATATAATAACAAGAG 2793  
DB 1789 AAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAACATATAATAACAAGAG 1848

QY 2794 TGCAGAAAGGTTTTTACTAGATTTTAGACATGACAGATATCAAGAGATAACCAAGCTAAAAAT 2853



Db 1849 |||||TGCGAAGAGT|||TACTAGAT|||TAGACATGCAGACTATCAAGATATAACCAAGCTAAAAAT 1308  
Qy 2854 |||||TAAGTGTGTGTCAAAAGAAAGCTGTCATTTAGGAAATGCTGTTCTTGGGAAACCCAAAGA 2913  
Db 1909 |||||TAAGTGTGTGTCAAAAGAAAGCTGTCATTTAGGAATGCTGTTCTTGGGAAACCCAAAGA 1968  
Qy 2914 |||||TACTGTGCTAGATGAACCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGATGGA 2973  
Db 1969 |||||TACTGTGCTAGATGAACCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGATGGA 2028  
Qy 2974 |||||ATCTTTTAAATACAGAAAAGCCAAATCGGTGACAGTGTTCAGTACHTCATTTTCATGGATG 3033  
Db 2029 |||||ATCTTTTAAATACAGAAAAGCCAAATCGGTGACAGTGTTCAGTACHTCATTTTCATGGATG 2088  
Qy 3034 |||||AAGCTGACATCTTTCAGATAGGAAGCTGTGTATCACAAAGGAATGCTGAAATGTGTTG 3093  
Db 2089 |||||AAGCTGACATCTTTCAGATAGGAAGCTGTGTATCACAAAGGAATGCTGAAATGTGTTG 2148  
Qy 3094 |||||GTTCTTCAATGTTCTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 3153  
Db 2149 |||||GTTCTTCAATGTTCTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2208  
Qy 3154 |||||ACAAATATTGCGCACAGAAATCTCTTCTTCACTGTTTAAACCAACATATATACCTGGAGCTA 3213  
Db 2209 |||||ACAAATATTGCGCACAGAAATCTCTTCTTCACTGTTTAAACCAACATATATACCTGGAGCTA 2268  
Qy 3214 |||||CTTTATTACAAAGATGACCAACACTGTGTATAGCTTGCCTTTCAAGGACATGGACA 3273  
Db 2269 |||||CTTTATTACAAAGATGACCAACACTGTGTATAGCTTGCCTTTCAAGGACATGGACA 2328  
Qy 3274 |||||AAATTTTCAGGTTGTTTCTGCCCCAGACAGTCAATCAAAATTTGGGTGTCAATTTCTTTATG 3333  
Db 2329 |||||AAATTTTCAGGTTGTTTCTGCCCCAGACAGTCAATCAAAATTTGGGTGTCAATTTCTTTATG 2388  
Qy 3334 |||||GTGTTTCCATGACGACTTTTGGGAAGACGTAATTTTAAAGCTAGAGTTGAAGCGAATTTG 3393  
Db 2389 |||||GTGTTTCCATGACGACTTTTGGGAAGACGTAATTTTAAAGCTAGAGTTGAAGCGAATTTG 2448  
Qy 3394 |||||ACCAAGCAGATATAGTGTATTTACTCAGCAGCCACTGGAGGAGAAATGGATTCAAAT 3453  
Db 2449 |||||ACCAAGCAGATATAGTGTATTTACTCAGCAGCCACTGGAGGAGAAATGGATTCAAAT 2508  
Qy 3454 |||||CTTTTGATGAATGGAAACAGAGCTTACTTATTTCTGAAACCAAGGCTTCTCTAGTGA 3513  
Db 2509 |||||CTTTTGATGAATGGAAACAGAGCTTACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGA 2568  
Qy 3514 |||||GCACCATGAGCTTTGGAAACCAACAGATGATACAAATAGCAAAAGTTTCATTTCTTACCT 3573  
Db 2569 |||||GCACCATGAGCTTTGGAAACCAACAGATGATACAAATAGCAAAAGTTTCATTTCTTACCT 2628  
Qy 3574 |||||TGAAACGTGAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 3633  
Db 2629 |||||TGAAACGTGAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 2688  
Qy 3634 |||||TTCAGATTTTATGTTTGGTTCATCACTCTTTTAAATATGCTGTGTTTCCCATCAAAC 3693  
Db 2689 |||||TTCAGATTTTATGTTTGGTTCATCACTCTTTTAAATATGCTGTGTTTCCCATCAAAC 2748  
Qy 3694 |||||TTGTTCCAGACTTATATTTTCTAAACCTGGAGACAAACCAATAAATACAAAACAAGTC 3753  
Db 2749 |||||TTGTTCCAGACTTATATTTTCTAAACCTGGAGACAAACCAATAAATACAAAACAAGTC 2808  
Qy 3754 |||||TGCTTCTTCAAAATCTGCTGACTCAGATATCAGTGTATCTATAGCTTTTTCACAGCC 3813  
Db 2809 |||||TGCTTCTTCAAAATCTGCTGACTCAGATATCAGTGTATCTATAGCTTTTTCACAGCC 2868  
Qy 3814 |||||AGAACATAATGCTGACGATGATTAATGACAGTGTATGATCGTGGCTCCCCATAGTG 3873  
Db 2869 |||||AGAACATAATGCTGACGATGATTAATGACAGTGTATGATCGTGGCTCCCCATAGTG 2928  
Qy 3874 |||||CGGCTTTTAAATGATGATTCATTCAGAAAAGGACTATGTTTTCAGAGCTGTTTTCACACAGTA 3933  
Db |||||CGGCTTTTAAATGATGATTCATTCAGAAAAGGACTATGTTTTCAGAGCTGTTTTCACACAGTA 4068

Db 2929 |||||CGGCTTTTAAATGATGATTCATTCAGAAAAGGACTATGTTTTCAGAGCTGTTTTCACACAGTA 2988  
Qy 3934 |||||CTATGTTTATTTCTTTACCTATATATAGTGAATATCATTTAGTAACTACTATCTTTTATCATTT 3993  
Db 2989 |||||CTATGTTTATTTCTTTTACCTATATATAGTGAATATCATTTAGTAACTACTATCTTTTATCATTT 3048  
Qy 3994 |||||TAAATGTGACTGAAACCAATCCAGATCTGGAGTACCCCATTTCTTTCAAGAAATTTACTGATA 4053  
Db 3049 |||||TAAATGTGACTGAAACCAATCCAGATCTGGAGTACCCCATTTCTTTCAAGAAATTTACTGATA 3108  
Qy 4054 |||||TAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTGTCTGGAATCATTTGTTACTGCAA 4113  
Db 3109 |||||TAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTGTCTGGAATCATTTGTTACTGCAA 3168  
Qy 4114 |||||TGCCACCTTACTTTTGCATGGAATAATGCAGAGAAATCATTAAGATCAAAAGCTTTATCTCAAC 4173  
Db 3169 |||||TGCCACCTTACTTTTGCATGGAATAATGCAGAGAAATCATTAAGATCAAAAGCTTTATCTCAAC 3228  
Qy 4174 |||||TTAAACCTTTTCAAGTCTTTTGCATCTGCATATTTGGATTTGGCAAGCTTTTGTGATATCC 4233  
Db 3229 |||||TTAAACCTTTTCAAGTCTTTTGCATCTGCATATTTGGATTTGGCAAGCTTTTGTGATATCC 3288  
Qy 4234 |||||CCTTATTTTATCATCTTATTTTTCATGCTAGGAACTTTATTTGGCAATTTTCATTTATGGAT 4293  
Db 3289 |||||CCTTATTTTATCATCTTATTTTTCATGCTAGGAACTTTATTTGGCAATTTTCATTTATGGAT 3348  
Qy 4294 |||||TATATTTTATCTGTAAGTCTTCTTGTCTGTGTTTTCCTTATTTGGCTTATTTGGTTTCCAT 4353  
Db 3349 |||||TATATTTTATCTGTAAGTCTTCTTGTCTGTGTTTTCCTTATTTGGCTTATTTGGTTTCCAT 3408  
Qy 4354 |||||CAGTTATTTCTGTTTCACTTATATTTTCTTCTTCACTTTTAAAGAAATTTTAAATACCAAAG 4413  
Db 3409 |||||CAGTTATTTCTGTTTCACTTATATTTTCTTCTTCACTTTTAAAGAAATTTTAAATACCAAAG 3468  
Qy 4414 |||||AAATTTGGTCAATTTATCTATTTCTGTGCGAGCGTGTGCTGTATTTGCAATCACTGAAATAA 4473  
Db 3469 |||||AAATTTGGTCAATTTATCTATTTCTGTGCGAGCGTGTGCTGTATTTGCAATCACTGAAATAA 3528  
Qy 4474 |||||CTTTCTTTTATGGGATACAAATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCATTC 4533  
Db 3529 |||||CTTTCTTTTATGGGATACAAATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCATTC 3588  
Qy 4534 |||||CAATCTATCCACTTCTAGTGTGCTGATTTCTTTTCAAAAGATTTCTTGAAGAATGTAC 4593  
Db 3589 |||||CAATCTATCCACTTCTAGTGTGCTGATTTCTTTTCAAAAGATTTCTTGAAGAATGTAC 3648  
Qy 4594 |||||GAAAAATGTGGACACCTATATCCATGGGATAGGCTTTTCAAGTGTGTTTATATTCGCTT 4653  
Db 3649 |||||GAAAAATGTGGACACCTATATCCATGGGATAGGCTTTTCAAGTGTGTTTATATTCGCTT 3708  
Qy 4654 |||||ACCTGAGTGTGCTGCTGATTTTCTTCAATATGAGAAAAATATGAGGCA 4713  
Db 3709 |||||ACCTGAGTGTGCTGCTGATTTTCTTCAATATGAGAAAAATATGAGGCA 3768  
Qy 4714 |||||GATCAATAAGAAAGATCCCTTTTTCAGAAACCTTTTCAAGAACTTAAATAAGGAGC 4773  
Db 3769 |||||GATCAATAAGAAAGATCCCTTTTTCAGAAACCTTTTCAAGAACTTAAATAAGGAGC 3828  
Qy 4774 |||||TTCAGAAACCAACAGCAATGAGGATGAAATGCAAGATGCAAGCTGAAAGACTAAAG 4833  
Db 3829 |||||TTCAGAAACCAACAGCAATGAGGATGAAATGCAAGATGCAAGCTGAAAGACTAAAG 3888  
Qy 4834 |||||TCAACAGCTGATGGGTGCTGAGTGTGTCAGGAGAAACCATCATTTATGCTGACCAAT 4893  
Db 3889 |||||TCAACAGCTGATGGGTGCTGAGTGTGTCAGGAGAAACCATCATTTATGCTGACCAAT 3948  
Qy 4894 |||||TGCTATAAGAAATATGATGCAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGG 4953  
Db 3949 |||||TGCTATAAGAAATATGATGCAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGG 4008  
Qy 4954 |||||CAACTAAATACATCTCTTTCTGTGTGTAAGAAAGGAGAGATCTTTAGGACTATTTGGTCCAA 5013  
Db 4009 |||||CAACTAAATACATCTCTTTCTGTGTGTAAGAAAGGAGAGATCTTTAGGACTATTTGGTCCAA 4068



```
QY 5014 ATGGTCTGCGCAAGACAAATATTAAATATCTGGTTGGTGATATTGAACCAACTTCAG 5073
Db 4069 ATGGTCTGCGCAAGACAAATATTAAATATCTGGTTGGTGATATTGAACCAACTTCAG 4128
QY 5074 GCCAGGTATTTTGGAGATATTCTCTTCAGAGACAAAGTGAAGATGATGATTCACCTGAAGT 5133
Db 4129 GCCAGATATTTTGGAGATATTCTCTTCAGAGTCAAGTGAAGATGATGATTCACCTGAAGT 4188
QY 5134 GTATGGTTACTGCTCTCAGATAAACCCTTTGTGGCCAGATATCTACATTCGAGGAACATT 5193
Db 4189 GTATGGTTACTGCTCTCAGATAAACCCTTTGTGGCCAGATATCTACATTCGAGGAACATT 4248
QY 5194 TTGAATTTATGAGCTGCTCAAGGATGAGTCAAGTGAAGTGAAGTCAAGTCAATAGTC 5253
Db 4249 TTGAATTTATGAGCTGCTCAAGGATGAGTCAAGTGAAGTGAAGTCAATAGTC 4308
QY 5254 GAATAACACATGCACTTGATTTAAAGAACATCTTCAGAGAGCTGTAAAGAACTACCTG 5313
Db 4309 GAATAACACATGCACTTGATTTAAAGAACATCTTCAGAGAGCTGTAAAGAACTACCTG 4368
QY 5314 CAGGAATCAAAACGAAAGTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTACTT 5373
Db 4369 CAGGAATCAAAACGAAAGTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTACTT 4428
QY 5374 TGCTAGTAGAACCATCTACAGGTATGATCCCAAGCCAAACAGCAGCATGTGGCGAGCAA 5433
Db 4429 TGCTAGTAGAACCATCTACAGGTATGATCCCAAGCCAAACAGCAGCATGTGGCGAGCAA 4488
QY 5434 TTCCGAATCGCATTTAAACACAGAAAGCGGGCTGCTATCTTCAGCACTCACTATATGAGG 5493
Db 4489 TTCCGAATCGCATTTAAACACAGAAAGCGGGCTGCTATCTTCAGCACTCACTATATGAGG 4548
QY 5494 AGGCAGAGGCTGCTGTGATCGAGTAGCTATCATGTTGCTGGGAGTTAAGATGATCG 5553
Db 4549 AGGCAGAGGCTGCTGTGATCGAGTAGCTATCATGTTGCTGGGAGTTAAGATGATCG 4608
QY 5554 GAACAGTACAACATCTAAAGAGTAAATTTGGAAAGCGCTACTTTTGGAAATTAATTTGA 5613
Db 4609 GAACAGTACAACATCTAAAGAGTAAATTTGGAAAGCGCTACTTTTGGAAATTAATTTGA 4668
QY 5614 AGGACTGGATAGAAAACCTTAGAAGTAGACCGCTTCAAGAGAGAAATTCAGTATATTTCC 5673
Db 4669 AGGACTGGATAGAAAACCTTAGAAGTAGACCGCTTCAAGAGAGAAATTCAGTATATTTCC 4728
QY 5674 CAAATCAAGCCGTCAGAAAGTTTTCCTCTATTTTGGCTTATAAAATTCCTAAGGAAG 5733
Db 4729 CAAATCAAGCCGTCAGAAAGTTTTCCTCTATTTTGGCTTATAAAATTCCTAAGGAAG 4788
QY 5734 ATGTTCACTCCCTTTCACAACTCTTTTAAAGCTGGAGAGCTTAACATGCTTTTCCCA 5793
Db 4789 ATGTTCACTCCCTTTCACAACTCTTTTAAAGCTGGAGAGCTTAACATGCTTTTCCCA 4848
QY 5794 TTGAAGATATAGCTTTCTCAAGCAACTTTGGAACAGGTTTGTAGAACTCACTAAAG 5853
Db 4849 TTGAAGATATAGCTTTCTCAAGCAACTTTGGAACAGGTTTGTAGAACTCACTAAAG 4908
QY 5854 AACAAAGAGGAGGAAGATAATAGTTGTGGAACATTTTAAACAGACACATTTGTGGGAACGAA 5913
Db 4909 AACAAAGAGGAGGAAGATAATAGTTGTGGAACATTTTAAACAGACACATTTGTGGGAACGAA 4968
QY 5914 CACAAGAGATAGAGTAGTATTTGAATTTGTATTTGCTCGCTTACTGGGACTTCT 5973
Db 4969 CACAAGAGATAGAGTAGTATTTGAATTTGTATTTGCTCGCTTACTGGGACTTCT 5028
QY 5974 TTCTTTTTCACCTTAATTTTAACTTTTGGTTTAAAGTTTATTTGGAATGGTAACTGGA 6033
Db 5029 TTCTTTTTCACCTTAATTTTAACTTTTGGTTTAAAGTTTATTTGGAATGGTAACTGGA 5088
QY 6034 GAACCAAG 6041
Db 5089 GAACCAAG 5096
```

```
RESULT 8
AX417824
LOCUS AX417824 4929 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 1 from Patent WO0231147.
ACCESSION AX417824
VERSION AX417824.1 GI:21522942
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Hu, Y. and Nepomichiv, B.
TITILE Human transporter proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0231147-A 1 18-APR-2002;
LEXICON GENETICS INC (US)
FEATURES
source
1..4929
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 1544 a 827 c 950 g 1604 t 4 others
ORIGIN
Query Match 75.4%; Score 4918.4; DB 6; Length 4929;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 4919; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY 1011 ATGTCACCTGCAATTAGGGAGGTAGGAGTTGGAGACAGACCCAGAACACTTCTTACTGAAG 1070
Db 1 ATGTCACCTGCAATTAGGGAGGTAGGAGTTGGAGACAGACCCAGAACACTTCTTACTGAAG 60
QY 1071 AATTACTTAATTAATGAGAACCAAAAAGAGTAGTGTTCAGGAATCTTTTCCACTA 1130
Db 61 AATTACTTAATTAATGAGAACCAAAAAGAGTAGTGTTCAGGAATCTTTTCCACTA 120
QY 1131 TTTTCTTTTATTTGGTTAATATTAAATAGCATGATGCATCCAAATAAGAAATATGAAGAA 1190
Db 121 TTTTCTTTTATTTGGTTAATATTAAATAGCATGATGCATCCAAATAAGAAATATGAAGAA 180
QY 1191 GTGCTTAATATAGAACTCAATCCTATGGAACAAGTTTACTCTTTTAAATCTAATCTTGA 1250
Db 181 GTGCTTAATATAGAACTCAATCCTATGGAACAAGTTTACTCTTTTAAATCTAATCTTGA 240
QY 1251 TATACTCCAGTCACTAATATTACAAGCAGCATCATGAGAAAGTGTCTACTCATCTA 1310
Db 241 TATACTCCAGTCACTAATATTACAAGCAGCATCATGAGAAAGTGTCTACTCATCTA 300
QY 1311 CCTGATGTCATTAATTTACTGAAGAAATATACAAATGAAAAGAAATGTTAAACATCCAGTCTC 1370
Db 301 CTTGATGTCATTAATTTACTGAAGAAATATACAAATGAAAAGAAATGTTAAACATCCAGTCTC 360
QY 1371 TCTAAGCCGAGCAACTTTGTAGGTGGTTTCAAAGACTCCCATGCTCTATGAACCTTCT 1430
Db 361 TCTAAGCCGAGCAACTTTGTAGGTGGTTTCAAAGACTCCCATGCTCTATGAACCTTCT 420
QY 1431 TTTTCTCTGATATGATTCAGTATCTTCTATTATTTATATGATTCAGAGCTGGCTGTTCA 1490
Db 421 TTTTCTCTGATATGATTCAGTATCTTCTATTATTTATGATTCAGAGCTGGCTGTTCA 480
QY 1491 AAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTTCACAGTTTTCACAGCATCCATA 1550
Db 481 AAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTTCACAGTTTTCACAGCATCCATA 540
QY 1551 GATGCTGCCATTTATACAGTTGAAGACCAATGTTTCTCTTTTGAAGAGCTGGAGTCAACT 1610
Db 541 GATGCTGCCATTTATACAGTTGAAGACCAATGTTTCTCTTTTGAAGAGCTGGAGTCAACT 600
QY 1611 AAAGCTGTTTATTTATGGGAGAACTGCTGTGTGTAGAAATAGATACCTTTCCCGAGGAGTA 1670
Db 601 AAAGCTGTTTATTTATGGGAGAACTGCTGTGTGTAGAAATAGATACCTTTCCCGAGGAGTA 660
```

Qy	1671	ATTTTAAATACCTAGTATAGCAATTTTCACCTTTTGGATACATTTTGGCAATTCATATC	1730
Db	661	ATTTTAAATACCTAGTATAGCAATTTTCACCTTTTGGATACATTTTGGCAATTCATATC	720
Qy	1731	GTAGCAGAAAAGAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACT	1790
Db	721	GTAGCAGAAAAGAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACT	780
Qy	1791	GCCTTTTGGCTTCTCTGGTCTCTCTATATACAAAGTTTAAATTTTCTATGTCCTCTTT	1850
Db	781	GCCTTTTGGCTTCTCTGGTCTCTCTATATACAAAGTTTAAATTTTCTATGTCCTCTTT	840
Qy	1851	ATGGCAGTCAATGCGACAGCTTCTTTGTTATTTCTCAAAAGTAGCAGCATTTGATATTT	1910
Db	841	ATGGCAGTCAATGCGACAGCTTCTTTGTTATTTCTCAAAAGTAGCAGCATTTGATATTT	900
Qy	1911	CTGCTTTTCTTTTATGATATCATCTGTATTTTCTTTTCTTTTAAATGCTGACACCTCTT	1970
Db	901	CTGCTTTTCTTTTATGATATCATCTGTATTTTCTTTTCTTTTAAATGCTGACACCTCTT	960
Qy	1971	TTTAAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTACTGTGGCTTTTGGATTT	2030
Db	961	TTTAAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTACTGTGGCTTTTGGATTT	1020
Qy	2031	ATTGGCCTTATGATAATCCTCATAGAAAGTTTCCCAAATCGTTAGTGTGGCTTTTCAGT	2090
Db	1021	ATTGGCCTTATGATAATCCTCATAGAAAGTTTCCCAAATCGTTAGTGTGGCTTTTCAGT	1080
Qy	2091	CCTTTCTGTCACGTACTTTTGTGATTTGGTATTTGCGACAGCTCATGATTTAGAAATTTT	2150
Db	1081	CCTTTCTGTCACGTACTTTTGTGATTTGGTATTTGCGACAGCTCATGATTTAGAAATTTT	1140
Qy	2151	AATGAAGTGCTTCATTTTCAAAATTTGACATGCGGCCCATATCTCTCTAAATTTATACAAT	2210
Db	1141	AATGAAGTGCTTCATTTTCAAAATTTGACATGCGGCCCATATCTCTCTAAATTTATACAAT	1200
Qy	2211	ATCATGCTCACATTAATAGTATATCTATGTCCTCTTGGCTGTCTATCTTTGATCAAGTC	2270
Db	1201	ATCATGCTCACATTAATAGTATATCTATGTCCTCTTGGCTGTCTATCTTTGATCAAGTC	1260
Qy	2271	ATTCCAGGGAAATTTGGCTTACGGAGATCATCTTTATTTTCTGAAGCCTTCATATGG	2330
Db	1261	ATTCCAGGGAAATTTGGCTTACGGAGATCATCTTTATTTTCTGAAGCCTTCATATGG	1320
Qy	2331	TCAAAGAGTAAAGAAATTTATGAGAGTTATCAGAGGGCAATGTTAATGAAATATTAGT	2390
Db	1321	TCAAAGAGTAAAGAAATTTATGAGAGTTATCAGAGGGCAATGTTAATGAAATATTAGT	1380
Qy	2391	TTTAGTGAAATTTATGAGCCAGTTTCTTCAGAAATTTGTAGGAAAAGAGCCATAAGAAAT	2450
Db	1381	TTTAGTGAAATTTATGAGCCAGTTTCTTCAGAAATTTGTAGGAAAAGAGCCATAAGAAAT	1440
Qy	2451	AGTGGTATTCAGAAAGACATACAGAAAGAGGGTGAATAATGTGGAGGCTTTGAGAAATTTG	2510
Db	1441	AGTGGTATTCAGAAAGACATACAGAAAGAGGGTGAATAATGTGGAGGCTTTGAGAAATTTG	1500
Qy	2511	TCATTTGACATATATGAGGGTCAGATTACTGCCCTTACTTGGCCACAGTGGAAAGAGGAAAG	2570
Db	1501	TCATTTGACATATATGAGGGTCAGATTACTGCCCTTACTTGGCCACAGTGGAAAGAGGAAAG	1560
Qy	2571	AGTACATTGATGAATATTTCTTTTGGGACTCTGCCACCTCTCTGATGGGTTTGGATCTATA	2630
Db	1561	AGTACATTGATGAATATTTCTTTTGGGACTCTGCCACCTCTCTGATGGGTTTGGATCTATA	1620
Qy	2631	TATGGACACAGAGTCTCAGAAATAGATGAATTTTGAAGCAGAAATAATGATTTGGCAAT	2690
Db	1621	TATGGACACAGAGTCTCAGAAATAGATGAATTTTGAAGCAGAAATAATGATTTGGCAAT	1680
Qy	2691	TGTCACAGTTAGATATACACTTTTGTATTTTGGACGTAGAGAAATTTTATCAATTTTG	2750
Db	1681	TGTCACAGTTAGATATACACTTTTGTATTTTGGACGTAGAGAAATTTTATCAATTTTG	1740
Qy	2751	GCCTTCAATCAAGGGATACAGCCCAACAATATAACAAGAGTGAGAGGTTTTTACTA	2810

Db	1741	GCCTTCAATCAAGGGATACAGCCCAATAATAACAAGAGTGAGAGGTTTTTACTA	1800
Qy	2811	GATTTAGACATGACAGACTATCAAGATAACCAAGCTAAAAAATTAAGTGTGTGCTCAAAA	2870
Db	1801	GATTTAGACATGACAGACTATCAAGATAACCAAGCTAAAAAATTAAGTGTGTGCTCAAAA	1860
Qy	2871	AGAAAGCTGTCAATTAGGAATTTGCTTTTGGGAACCCAAAGATACCTGCTGTAGATGAA	2930
Db	1861	AGAAAGCTGTCAATTAGGAATTTGCTTTTGGGAACCCAAAGATACCTGCTGTAGATGAA	1920
Qy	2931	CMACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGTATGTAATCTTTTAAATACAGA	2990
Db	1921	CMACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGTATGTAATCTTTTAAATACAGA	1980
Qy	2991	AAAGCCAATCGGCTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATTTCTTGA	3050
Db	1981	AAAGCCAATCGGCTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATTTCTTGA	2040
Qy	3051	GATAGAAAAGCTGTGATATCACAAAGGAATGCTGAAATGTGTGGTTCTTCAATGTTCCCTC	3110
Db	2041	GATAGAAAAGCTGTGATATCACAAAGGAATGCTGAAATGTGTGGTTCTTCAATGTTCCCTC	2100
Qy	3111	AAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAAATATTGTGCCACA	3170
Db	2101	AAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAAATATTGTGCCACA	2160
Qy	3171	GAACTCTTTCTTCTACCTGTTAAACAACATATACCTGGAGCTACTTTTATACACAGAAT	3230
Db	2161	GAACTCTTTCTTCTACCTGTTAAACAACATATACCTGGAGCTACTTTTATACACAGAAT	2220
Qy	3231	GACCAACAACTGTGTATAGCTTGCCTTTCAAGGACATGACAAAATTTTTCAGGTTTGT	3290
Db	2221	GACCAACAACTGTGTATAGCTTGCCTTTCAAGGACATGACAAAATTTTTCAGGTTTGT	2280
Qy	3291	TCGCCCCATAGACAGTCAATTCAAATTTGGGTGTCAATTTCTTATGGTGTTCATGACGACT	3350
Db	2281	TCGCCCCATAGACAGTCAATTCAAATTTGGGTGTCAATTTCTTATGGGTGTTCATGACGACT	2340
Qy	3351	TTGGAGAGAGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATATTAGT	3410
Db	2341	TTGGAGAGAGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATATTAGT	2400
Qy	3411	GTATTTACTCAGCAGCCACTGGAGGAGAAATGGATTCAAAATCTTTTGAATGAATGGAA	3470
Db	2401	GTATTTACTCAGCAGCCACTGGAGGAGAAATGGATTCAAAATCTTTTGAATGAATGGAA	2460
Qy	3471	CAGAGCTTACTTATTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTGG	3530
Db	2461	CAGAGCTTACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTGG	2520
Qy	3531	AAACAACAGATGTATACAATAGCAAAAGTTTCATTTCTTTTACCTTGAACCTGAAAGTAA	3590
Db	2521	AAACAACAGATGTATACAATAGCAAAAGTTTCATTTCTTTTACCTTGAACCTGAAAGTAA	2580
Qy	3591	TCAGTGAGATCAGTGTGTCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTT	3650
Db	2581	TCAGTGAGATCAGTGTGTCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTT	2640
Qy	3651	TTGGTTTCACTCCTTTTAAAAATGCTGTGGTTTCCCATCAAACTGTTCAGACTTATAT	3710
Db	2641	TTGGTTTCACTCCTTTTAAAAATGCTGTGGTTTCCCATCAAACTGTTCAGACTTATAT	2700
Qy	3711	TTTCTAAAAACCTGGAGACAAACCAATATAAACAAGCTGCTCTTCTTCAAAATTTCT	3770
Db	2701	TTTCTAAAAACCTGGAGACAAACCAATATAAACAAGCTGCTCTTCTTCAAAATTTCT	2760
Qy	3771	GCTGACTCAGATATCAGTGTATTTAGCTTTTCAAGCCAGAACATATAATGGTGACG	3830
Db	2761	GCTGACTCAGATATCAGTGTATTTAGCTTTTCAAGCCAGAACATATAATGGTGACG	2820
Qy	3831	ATGATTAATGACAGTACTATGATCGGTGCTCCCATAGTGGGCTTTTAAATGTGATG	3890

2821	ATGATTAAATGACAGTGA	CTATGTATTC	CGTGGCTCCCACTAGT	CGGGCTTTAAATGTGRTG	2880	
Qy	CATTACAAAAGGACTATG	TTTTTTGGACGTGTTTT	CAACAGTACTATG	GGTTTTATTTCTTTA	3950	
	2881	CATTACAAAAGGACTATG	TTTTTTGGACGTGTTTT	CAACAGTACTATG	GGTTTTATTTCTTTA	2940
3951	CCTATATTTAGTGAATAT	CAATTTAGTAACTACT	ATCTATCTTTTATCAT	TTTAAATGTGACTGAAACC	4010	
Db	CCTATATTTAGTGAATAT	CAATTTAGTAACTACT	ATCTATCTTTTATCAT	TTTAAATGTGACTGAAACC	3000	
	4011	ATCCAGATCTGGAGTAC	CCCCATCTCTTTCAGAAAAAT	TACTGATATAGTTTTTAAATTTGAG	4070	
3001	ATCCAGATCTGGAGTAC	CCCCATCTCTTTCAGAAAAAT	TACTGATATAGTTTTTAAATTTGAG	3060		
Qy	CTGTATTTTTCAAGCAG	CTTTTGCTTTGGAAATCA	TATGTTTACTGCAAT	TGCCACTTACTTTTGCC	4130	
	3061	CTGTATTTTTCAAGCAG	CTTTTGCTTTGGAAATCA	TATGTTTACTGCAAT	TGCCACTTACTTTTGCC	3120
4131	ATGGAAAAATGCAGAAAT	CAATAAGATCAAAAGCT	TATATCTCAACTTAAAC	TTTCAGTCTT	4190	
Db	ATGGAAAAATGCAGAAAT	CAATAAGATCAAAAGCT	TATATCTCAACTTAAAC	TTTCAGTCTT	3180	
	4191	TTGCCATCTGCATAT	TGGATTTGGCAAGCTG	TGTTGATATCCCTTATTTTTTATCAT	TT	4250
3181	TTGCCATCTGCATAT	TGGATTTGGCAAGCTG	TGTTGATATCCCTTATTTTTTATCAT	TT	3240	
Qy	CTTATTTTGTATGCT	PAGAAGCTTACTCGCA	TTTCACTATGGAAT	TATATTTTTATAC	4310	
	3241	CTTATTTTGTATGCT	PAGAAGCTTACTCGCA	TTTCACTATGGAAT	TATATTTTTATAC	3300
4311	AAGTTCCTTGCTGTG	GGTTTTTTTGGCTTTAT	TGGTTTATGTTCCAT	CAGTTATTTCTGTTCACT	4370	
Db	AAGTTCCTTGCTGTG	GGTTTTTTTGGCTTTAT	TGGTTTATGTTCCAT	CAGTTATTTCTGTTCACT	3360	
	4371	TATATGCTTCTTTTCA	CTTTTAAAGAAATTTT	TAAATACCAAGAA	TTTTTGGTCA	4430
3361	TATATGCTTCTTTTCA	CTTTTAAAGAAATTTT	TAAATACCAAGAA	TTTTTGGTCA	3420	
Qy	TATTTCTGTGGCAGG	TGTTGNCCTGTAT	TGCAATCACTGAA	ATACTTTCTTTATG	4490	
	3421	TATTTCTGTGGCAGG	TGTTGNCCTGTAT	TGCAATCACTGAA	ATACTTTCTTTATG	3480
4491	ACAAATGCAACTAT	TTCTTCAATTAATG	TCCTTTTGTATCAT	CACTTCCAATCTTCA	4550	
Db	ACAAATGCAACTAT	TTCTTCAATTAATG	TCCTTTTGTATCAT	CACTTCCAATCTTCA	3540	
	4551	GGTTGCCCTGATTTCT	TTTTCATAAAGATTT	CTTGGAGAATGTAC	AGAAAAATGTGGACACC	4610
3541	GGTTGCCCTGATTTCT	TTTTCATAAAGATTT	CTTGGAGAATGTAC	AGAAAAATGTGGACACC	3600	
Qy	TATAATCCATGGGAT	GAGCTTTACGATAGCT	TGTTATATCGCTT	TACCTGCACTG	4670	
	3601	TATAATCCATGGGAT	GAGCTTTACGATAGCT	TGTTATATCGCTT	TACCTGCACTG	3660
4671	TGGATTTTCTCTTTA	CAATACTATGAGAAAAA	TATGGAGCGAGAT	CAATAAGAAAAAGAT	4730	
Db	TGGATTTTCTCTTTA	CAATACTATGAGAAAAA	TATGGAGCGAGAT	CAATAAGAAAAAGAT	3720	
	4731	CCCTTTTTTCAGAAAC	CTTTTCAACGAGTCT	TAAAAATAGGAAGCT	TTCAGAAACCAACAGAC	4790
3721	CCCTTTTTTCAGAAAC	CTTTTCAACGAGTCT	TAAAAATAGGAAGCT	TTCAGAAACCAACAGAC	3780	
Qy	AATCAGGATGAAGAT	GAAGATGTCAAGAGCT	GAAGAGCTTAAAGG	TCAAGAGCTGATGGGT	4850	
	3781	AATCAGGATGAAGAT	GAAGATGTCAAGAGCT	GAAGAGCTTAAAGG	TCAAGAGCTGATGGGT	3840
4851	TGCCAGTGTGTTGAG	GAGAAACCAATCCAT	TATGTGTGAGCAAT	TTTGCATAAAGAAATATGAT	4910	
Db	TGCCAGTGTGTTGAG	GAGAAACCAATCCAT	TATGTGTGAGCAAT	TTTGCATAAAGAAATATGAT	3900	
	4911	GACAAGAAAGATTTTCT	TTTCTTTTCAAGAAAAAG	TAAAGAAAGTGGCA	CTATAATCATCTCT	4970
3901	GACAAGAAAGATTTTCT	TTTCTTTTCAAGAAAAAG	TAAAGAAAGTGGCA	CTATAATCATCTCT	3960	

[illegible]



1801 DB GATTTAGACATGCGAGCTATCAAGATAACCAAGCTAAAAAATAAGTGGTGTCAAAA 1860  
2871 QY AGAAGCTGTCATTTAGGAATGCTGTTCTGGGAACCAAGATACCTGCTGCTAGTGA 2930  
1861 DB AGAAGCTGTCATTTAGGAATGCTGTTCTGGGAACCAAGATACCTGCTGCTAGTGA 1920  
2931 QY CCAACAGCTGAATGGACCCCTGTTCTGCAATATTTGTAATCTTTTAAATACAGA 2990  
1921 DB CCAACAGCTGAATGGACCCCTGTTCTGCAATATTTGTAATCTTTTAAATACAGA 1980  
2991 QY AAAGCCAAATCGGCTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATTTCTGCA 3050  
1981 DB AAAGCCAAATCGGCTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATTTCTGCA 2040  
3051 QY GATAGGAAGCTGTGATATCAAGGAATGCTGAATGTTGGTCTTCAATGTTCCCTC 3110  
2041 DB GATAGGAAGCTGTGATATCAAGGAATGCTGAATGTTGGTCTTCAATGTTCCCTC 2100  
3111 QY AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTCGCCACA 3170  
2101 DB AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTCGCCACA 2160  
3171 QY GAATCTCTTCTTCACTGTTTAAACACATATACCTGGAGCTACTTTTATTAACAAGAA 3230  
2161 DB GAATCTCTTCTTCACTGTTTAAACACATATACCTGGAGCTACTTTTATTAACAAGAA 2220  
3231 QY GACCAACAACTGTGTATAGTGTGCTTCAAGGACATGACAAATTTTCAGGTTGTTT 3290  
2221 DB GACCAACAACTGTGTATAGTGTGCTTCAAGGACATGACAAATTTTCAGGTTGTTT 2280  
3291 QY TCTGCCCTAGACAGTCAATCAAAATTTGGGTGTCATTTCTTATGTTTCCATGAGCACT 3350  
2281 DB TCTGCCCTAGACAGTCAATCAAAATTTGGGTGTCATTTCTTATGTTTCCATGAGCACT 2340  
3351 QY TTGGAAGACGCTATTTTAAAGCTAGAAAGTTGAAGCAGAAATTTGACCAAGCAGATTAAGT 3410  
2341 DB TTGGAAGACGCTATTTTAAAGCTAGAAAGTTGAAGCAGAAATTTGACCAAGCAGATTAAGT 2400  
3411 QY GTATTTACTGAGCAGCTGAGGAGAAATGGAATCAAAATCTTTTGTATGAATGGAA 3470  
2401 DB GTATTTACTGAGCAGCTGAGGAGAAATGGAATCAAAATCTTTTGTATGAATGGAA 2460  
3471 QY CAGAGCTTACTTATTTCTGGAACCAAGCTTCTCTAGTGAGCACCATGAGCCTTTGG 3530  
2461 DB CAGAGCTTACTTATTTCTGGAACCAAGCTTCTCTAGTGAGCACCATGAGCCTTTGG 2520  
3531 QY AAAACAACAGATGTATACAATAGCAAAAGTTTCATTTCTTTTACCTTGAACCGTGAAGTAA 3590  
2521 DB AAAACAACAGATGTATACAATAGCAAAAGTTTCATTTCTTTTACCTTGAACCGTGAAGTAA 2580  
3591 QY TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTTCAGATTTTATGTTT 3650  
2581 DB TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTTCAGATTTTATGTTT 2640  
3651 QY TTGCTTCACTCACTTTTAAATGCTGTTGTTCCCATCAAACTGCTTCTTCAAAATTCCT 3710  
2641 DB TTGCTTCACTCACTTTTAAATGCTGTTGTTCCCATCAAACTGCTTCTTCAAAATTCCT 2700  
3711 QY TTTCTTAAACCTGGAGACAAACCAATATAAATAAACAAGTCTGCTTCTTCAAAATTCCT 3770  
2701 DB TTTCTTAAACCTGGAGACAAACCAATATAAATAAACAAGTCTGCTTCTTCAAAATTCCT 2760  
3771 QY GCT-----GACTCAGATATCAGTGTATTTAGCTTTTTCACAAAGC 3812  
2761 DB GCTGGTGAGATGTGTTNGTGAGACTCAGATATCAGTGTATTTAGCTTTTTCACAAAGC 2820  
3813 QY CAGAACATATGTTGACGATTAATGACAGTACTATGATCCGTTGCTCCCAATAGT 3872  
2821 DB CAGAACATATGTTGACGATTAATGACAGTACTATGATCCGTTGCTCCCAATAGT 2880  
3873 QY GCGGCTTTTAAATGTGATGCAATTCAGAAAAGGACTATGTTTTCAGCAGTGTTTTCAACAGT 3932

2881 DB GCGGCTTTTAAATGTGATGCAATTCAGAAAAGGACTATGTTTTCAGAGCTGTTTTCACAGT 2940  
3933 QY ACTATGTTTATTTCTTTTACCTATATTTAGTGAATATCATTTAGTAACTACTATCTTTATCAT 3992  
2941 DB ACTATGTTTATTTCTTTTACCTATATTTAGTGAATATCATTTAGTAACTACTATCTTTATCAT 3000  
3993 QY TTTAAATGTGACTGAAACCAATCCAGATCTGGAGTACCCATTTCTTTCAAGAAATTAAGTAT 4052  
3001 DB TTTAAATGTGACTGAAACCAATCCAGATCTGGAGTACCCATTTCTTTCAAGAAATTAAGTAT 3060  
4053 QY ATAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTGTCTGG-AAATCATTTGTACTGCG 4111  
3061 DB ATAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTGTCTGGAAATCATTTGTAACTGCG 3120  
4112 QY AATGCCACCTTACTTTTGGCATGGAATGCAAGAAATCATAGATCAAAAGCTTATATACTCA 4171  
3121 DB AATGCCACCTTACTTTTGGCATGGAATGCAAGAAATCATAGATCAAAAGCTTATATACTCA 3161  
4172 QY ACTTAAACCTTTTCAGGCTTTTGGCCATCTGTCATATTTGGATTTGGACAAAGCTGTTTGTATAT 4231  
3162 DB -----AGGTCITTTTGGCCATCTGTCATATTTGGATTTGGACAAAGCTGTTTGTATAT 3209  
4232 QY CCCCTTATTTTATATCTTTTATTTTGGATCTGATGCTAGGAAGCTTACTGGCATTTTCAATATGG 4291  
3210 DB CCCCTTATTTTATATCTTTTATTTTGGATCTGATGCTAGGAAGCTTATTTGGCATTTTCAATATGG 3269  
4292 QY ATTATATTTTATATCTGTAAGTTCCTGCTGCTGTTTTCGCTTATTTGGTATATGTTTCC 4351  
3270 DB ATTATATTTTATATCTGTAAGTTCCTGCTGCTGTTTTCGCTTATTTGGTATATGTTTCC 3329  
4352 QY ATCAGTATTTCTGTTTCACTTATATTTCTTTTCACTTTTAAAGAAATTTTAAATACCAA 4411  
3330 DB ATCAGTATTTCTGTTTCACTTATATTTCTTTTCACTTTTAAAGAAATTTTAAATACCAA 3389  
4412 QY AGAATTTTGGTCAATTTATCTATTTCTGTCGACGCTTGTCTGATTTGCAATCACTGAAAT 4471  
3390 DB AGAATTTTGGTCAATTTATCTATTTCTGTCGACGCTTGTCTGATTTGCAATCACTGAAAT 3449  
4472 QY AACTTTCTTTATGGGATACACAATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCAT 4531  
3450 DB AACTTTCTTTATGGGATACACAATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCAT 3509  
4532 QY TCCAATCTATCCACTTTAGGTTGCTGATTTCTTTTCAATAAGATTTCTTGAAGAATGT 4591  
3510 DB TCCAATCTATCCACTTTAGGTTGCTGATTTCTTTTCAATAAGATTTCTTGAAGAATGT 3569  
4592 QY ACGAAAAATGTGGACACCTTAAATCCATGGATAGGCTTTTCAGTAGCTTGTATATCGCC 4651  
3570 DB ACGAAAAATGTGGACACCTTAAATCCATGGATAGGCTTTTCAGTAGCTTGTATATCGCC 3629  
4652 QY TTACCTGCAAGTGTGCTGATTTTCTCTTCAATACTATGAGAAAAATATGAGG 4711  
3630 DB TTACCTGCAAGTGTGCTGATTTTCTCTTCAATACTATGAGAAAAATATGAGG 3689  
4712 QY CAGATCAATAAGAAAGATFCCCTTTTTCAGAAACCTTTTCAACGAAAGTCTTAAATAATAGGAA 4771  
3690 DB CAGATCAATAAGAAAGATFCCCTTTTTCAGAAACCTTTTCAACGAAAGTCTTAAATAATAGGAA 3749  
4772 QY GCTTCAGAACCAACAGCAATGAGATGAAGATGAAGATGCAAGCTGGAAGACTAA 4831  
3750 DB GCTTCAGAACCAACAGCAATGAGATGAAGATGAAGATGCAAGCTGGAAGACTAA 3809  
4832 QY GGTCAAGAGCTGATGGGTTGCGAGTGTGTCAGGAGAAACCATTCATTTATCGTCAGCAA 4891  
3810 DB GGTCAAGAGCTGATGGGTTGCGAGTGTGTCAGGAGAAACCATTCATTTATGTCAGCAA 3869  
4892 QY TTTTGCATAAAGATATGATGACAAGAAATTTTCTTTTCAAGAAAAAGTAAAGAAAGT 4951  
3870 DB TTTTGCATAAAGATATGATGACAAGAAATTTTCTTTTCAAGAAAAAGTAAAGAAAGT 3929  
4952 QY GCGCAACTAAATACATCTTTCTGTTGTAAGAAAGGAGAGATCTTAGGACTATTTGGGCTCC 5011  
3930 DB GCGCAACTAAATACATCTTTCTGTTGTAAGAAAGGAGAGATCTTAGGACTATTTGGGCTCC 3989





Qy	1791	GCCTTTGGCTTCTCCTGGGTTCTTCTATATACAAAGTTTAATTTTCTTATATGTCCTTCTTT	1850
Db	781	GCCTTTGGCTTCTCCTGGGTTCTTCTATAWACAAAGTTTAATTTTCTTATATGTCCTTCTTT	840
Qy	1851	ATGCGAGTCATTCGGACAGCTTCTTTGTTATTTTTCCTCAAAGTAGCAGCATTTGTATATTT	1910
Db	841	ATGCGAGTCATTCGGACAGCTTCTTTGTTATTTTTCCTCAAAGTAGCAGCATTTGTATATTT	900
Qy	1911	CTGCTTTTTTTCCTTTATGGAATTATCATCTGTATATTTTTCCTTTAATTCGTGATATTT	1970
Db	901	CTGCTTTTTTTCCTTTATGGAATTATCATCTGTATATTTTTCCTTTAATTCGTGATATTT	960
Qy	1971	TTTAAAAATCAAACAATGTCGGAATAGTTGAAATTTTTTGTATCTGTGGCTTTGGATTT	2030
Db	961	TTTAAAAATCAAACAATGTCGGAATAGTTGAAATTTTTTGTATCTGTGGCTTTGGATTT	1020
Qy	2031	ATTGGCCTTATGATATATCCTCATAGAAAGTTTTTCCAAATCGTTAGTGTGGCTTTTCAGT	2090
Db	1021	ATTGGCCTTATGATATATCCTCATAGAAAGTTTTTCCAAATCGTTAGTGTGGCTTTTCAGT	1080
Qy	2091	CCTTTCTGTCACATGTAATCTTTTGATTGGTATTTGCAACAGTCATGCAATTTAGAAATTTT	2150
Db	1081	CCTTTCTGTCACATGTAATCTTTTGATTGGTATTTGCAACAGTCATGCAATTTAGAAATTTT	1140
Qy	2151	AATGAAGGTGCTTCATTTTCAAAATTTGATCGAGGCCATATCCTCTAAATTAACAATTT	2210
Db	1141	AATGAAGGTGCTTCATTTTCAAAATTTGATCGAGGCCATATCCTCTAAATTAACAATTT	1200
Qy	2211	ATCATGCTCACACTTAATAGTATATTTCTATGTCCTCTTGGCTGTCTATCTTTGATCAAGTC	2270
Db	1201	ATCATGCTCACACTTAATAGTATATTTCTATGTCCTCTTGGCTGTCTATCTTTGATCAAGTC	1260
Qy	2271	ATTCCAGGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGGAAGCCTTCATATGG	2330
Db	1261	ATTCCAGGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGGAAGCCTTCATATGG	1320
Qy	2331	TCAAAGAGTAAAAAGAAATTAATGAGGAGTTATCAGAGGGCAATGTTAATGAAATATTAGT	2390
Db	1321	TCAAAGAGTAAAAAGAAATTAATGAGGAGTTATCAGAGGGCAATGTTAATGAAATATTAGT	1380
Qy	2391	TTTATGTAATTTATGAGCCAGTTTCTTCAGAAATTTGTAGAAAAGNAGCCATAAGAAATTT	2450
Db	1381	TTTATGTAATTTATGAGCCAGTTTCTTCAGAAATTTGTAGAAAAGNAGCCATAAGAAATTT	1440
Qy	2451	AGTGGTATTCAGAAAGACATACAGAAAGAGGGTGAAATAGTGGAGGCTTTTGAGAAATTTG	2510
Db	1441	AGTGGTATTCAGAAAGACATACAGAAAGAGGGTGAAATAGTGGAGGCTTTTGAGAAATTTG	1500
Qy	2511	TCATTTGACATATATGAGGGTCAGATTACTGCCCTTACTTTGCCACAGTGGAAACAGAAAG	2570
Db	1501	TCATTTGACATATATGAGGGTCAGATTACTGCCCTTACTTTGCCACAGTGGAAACAGAAAG	1560
Qy	2571	AGTACATTTGATGAATATTTCTTTTGGGACTCTGCCACCTTCTGATGGGTTTGATCTATA	2630
Db	1561	AGTACATTTGATGAATATTTCTTTTGGGACTCTGCCACCTTCTGATGGGTTTGATCTATA	1620
Qy	2631	TATGGACACAGAGTCCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAAATGATTTGGCATT	2690
Db	1621	TATGGACACAGAGTCCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAAATGATTTGGCATT	1680
Qy	2691	TGTCCACAGTTAGATATACACTTTTGATGTTTTTGACAGTAGAAGAAAAATTTTCAATTTTG	2750
Db	1681	TGTCCACAGTTAGATATACACTTTTGATGTTTTTGACAGTAGAAGAAAAATTTTCAATTTTG	1740
Qy	2751	GCTTCAATCAAAGGGATACAGCCAAACAATATAATAACAAGATGCGAGAAGTTTACTATA	2810
Db	1741	GCTTCAATCAAAGGGATACAGCCAAACAATATAATAACAAGATGCGAGAAGTTTACTATA	1800
Qy	2811	GATTTAGACATGCGAGTCTCAAGAGATAACCAAGCTAAAAAATTAAGTGTGTGTCAAAAA	2870
Db	1801	GATTTAGACATGCGAGTCTCAAGAGATAACCAAGCTAAAAAATTAAGTGTGTGTCAAAAA	1860
Qy	2871	AGAAAGCTGTCTATTAGGAAATTTGCTGTTCTTTGGGAACCAAGATACCTGCTGCTAGATGAA	2930

1861	AGAAAGCTGTCA	TTTAGGAATTC	GTGTTCTT	GGGAACCC	AAAGATACT	GTCTGCTAGATGAA	1921
2931	CCAA	CAGCTGGAAT	TGGACCCCTG	TCTCGACATA	TTGTATGSAAT	CTTTTAAAAATACAGA	2990
1921	CCAA	CAGCTGGAAT	TGGACCCCTG	TCTCGACATA	TTGTATGSAAT	CTTTTAAAAATACAGA	1980
2991	AAAG	CCAAATCGGGTGA	CAGTGTTC	TAGTACTCA	TTTCTATGAGATGAAG	CTGACATTTCTTGCA	3050
1981	AAAG	CCAAATCGGGTGA	CAGTGTTC	TAGTACTCA	TTTCTATGAGATGAAG	CTGACATTTCTTGCA	2040
3051	GATAGGAAG	CTGTGATATCAC	AAGGAATGCT	GAAATGT	TTGTTGTTCTT	CAATGTTCTCTC	3110
2041	GATAGGAAG	CTGTGATATCAC	AAGGAATGCT	GAAATGT	TTGTTGTTCTT	CAATGTTCTCTC	2100
3111	AAAA	GTAATCGGGGATCG	CGCTAC	CGCCTGAGCATGT	TACATAGACAAA	TATTTGTGCCACA	3170
2101	AAAA	GTAATCGGGGATCG	CGCTAC	CGCCTGAGCATGT	TACATAGACAAA	TATTTGTGCCACA	2160
3171	GAAT	CTCTTTCTT	CAC	TGGTTAAACAACATATA	CTCTGGAGCTAT	CTTTATATACACAGAA	3230
2161	GAAT	CTCTTTCTT	CAC	TGGTTAAACAACATATA	CTCTGGAGCTAT	CTTTATATACACAGAA	2220
3231	GACCAAC	ACTTGTGTATAG	CTTGCCCTTT	CAAGGACATG	GACAAATTT	CTCAGGTTCGTTT	3290
2221	GACCAAC	ACTTGTGTATAG	CTTGCCCTTT	CAAGGACATG	GACAAATTT	CTCAGGTTCGTTT	2280
3291	TC	TGCCCTTAGACAG	CTCAAT	CAAAATTTGGGGTGT	CAATTTCTTATGTGTT	TTTCCATGACGACT	3350
2281	TC	TGCCCTTAGACAG	CTCAAT	CAAAATTTGGGGTGT	CAATTTCTTATGTGTT	TTTCCATGACGACT	2340
3351	TTGGAAG	AGCTGATATTTT	TAAAGCTAG	AAGTTGAAG	CGAGAAATTTGA	CCAAAGCAGATATATAGT	3410
2341	TTGGAAG	AGCTGATATTTT	TAAAGCTAG	AAGTTGAAG	CGAGAAATTTGA	CCAAAGCAGATATATAGT	2400
3411	GTA	TTTACTGACGAC	CACCTGG	AGGAAGAAATGGAT	TTCAAAATCTTTT	GATGAAATCGAA	3470
2401	GTA	TTTACTGACGAC	CACCTGG	AGGAAGAAATGGAT	TTCAAAATCTTTT	GATGAAATCGAA	2460
3471	CAGAG	CTTACTATTTCT	TTTCTG	TAAACCAAGGCTT	CTCTAGTGAGCACC	ATGAGCCTTTTG	3530
2461	CAGAG	CTTACTATTTCT	TTTCTG	TAAACCAAGGCTT	CTCTAGTGAGCACC	ATGAGCCTTTTG	2520
3531	AAACAAC	AGATGTATACAA	TAGCAAAAG	TTTCATTTCTT	TAAACCTGTAAC	CGTGAAGTAAA	3590
2521	AAACAAC	AGATGTATACAA	TAGCAAAAG	TTTCATTTCTT	TAAACCTGTAAC	CGTGAAGTAAA	2580
3591	TCAG	TGAGATCAGTGT	TGCTTCTG	CTTTTAAATTTT	TTTTCACAGTTC	CAGATTTTATGTTT	3650
2581	TCAG	TGAGATCAGTGT	TGCTTCTG	CTTTTAAATTTT	TTTTCACAGTTC	CAGATTTTATGTTT	2640
3651	TTG	TTTCATCTCTTTT	TAAAAATG	CTGTGTTTCC	CAATCAAACTTTG	TTTCCAGACTTATAT	3710
2641	TTG	TTTCATCTCTTTT	TAAAAATG	CTGTGTTTCC	CAATCAAACTTTG	TTTCCAGACTTATAT	2700
3711	TTTT	CTAAAA	CCTGGAGACAA	CCACAT	AAATACAAAC	CAAGTCTGCTTCTT	3770
2701	TTTT	CTAAAA	CCTGGAGACAA	CCACAT	AAATACAAAC	CAAGTCTGCTTCTT	2760
3771	GCTG	ATCAGATATCAG	TGATCTTA	TTAGCTTTT	TTCACAGCCAGAA	CATAATGGTGACG	3830
2761	GCTG	ATCAGATATCAG	TGATCTTA	TTAGCTTTT	TTCACAGCCAGAA	CATAATGGTGACG	2820
3831	ATG	GATTAAATGAC	AGTACTAT	GTATCCG	TGGCTCCCATAGT	CGCGCTTTAAATGTGATG	3890
2821	ATG	GATTAAATGAC	AGTACTAT	GTATCCG	TGGCTCCCATAGT	CGCGCTTTAAATGTGATG	2880
3891	CATT	CAGAAAAGGACT	ATGTTTTG	CAGCTGTTTT	CAACAGTACT	TATGTTTCTTTA	3950
2881	CATT	CAGAAAAGGACT	ATGTTTTG	CAGCTGTTTT	CAACAGTACT	TATGTTTCTTTA	2940
3951	CCTA	TATTAGTGAA	TATCAT	TTAGTAACT	ACTATCTTTT	ATCATTTAAATGTGACTGAAACC	4010



Db 2941 CCTATTAGTGAATATCATTAGTAACATACTATCTTTATCAATTTAAATGTGACTGAAC 3000  
Qy 4011 ATCCAGATCTGGAGTACCCCAATCTTTCAAGAAATTAAGTATAGTTTTTAAATTTGAG 4070  
Db 3001 ATCCAGATCTGGAGTACCCCAATCTTTCAAGAAATTAAGTATAGTTTTTAAATTTGAG 3060  
Qy 4071 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCAACCTTTACTTTGCC 4130  
Db 3061 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCAACCTTTACTTTGCC 3120  
Qy 4131 ATGGAAATCCAGAGATCATAGATCAAGAGCTTTATCTCAACTTAACTTTTCAAGTCTTT 4190  
Db 3121 ATGGAAATCCAGAGATCATAGATCAAGAGCTTTATCTCAACTTAACTTTTCAAGTCTTT 3180  
Qy 4191 TTGCCATCTGCATATTTGGATTGGACAGCTTTGTTGATATCCCTTTATTTTTTATCAAT 4250  
Db 3181 TTGCCATCTGCATATTTGGATTGGACAGCTTTGTTGATATCCCTTTATTTTTATCAAT 3240  
Qy 4251 CTTATTTTGATGCTAGGAAGCTTACTGGCAATTTCAATATGGAATATATTTTTTATCTGTA 4310  
Db 3241 CTTATTTTGATGCTAGGAAGCTTACTGGCAATTTCAATATGGAATATATTTTTTATCTGTA 3300  
Qy 4311 AAGTTCCTTCTGCTGCTTTTGGCTTATTTGGTTATGTTCCATCAGTTATCTGTTCACT 4370  
Db 3301 AAGTTCCTTCTGCTGCTTTTGGCTTATTTGGTTATGTTCCATCAGTTATCTGTTCACT 3360  
Qy 4371 TATATTGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 4430  
Db 3361 TATATTGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3420  
Qy 4431 TATTTCTGTCAGCGTTGCTTGTATTGCAATCACTGAAATAAATTTCTTTTATGGGATAC 4490  
Db 3421 TATTTCTGTCAGCGTTGCTTGTATTGCAATCACTGAAATAAATTTCTTTTATGGGATAC 3480  
Qy 4491 ACAATGCAACTATTCTTCAATATGCTTTTGTATGATCATCTCAATCTATCACTTTCTA 4550  
Db 3481 ACAATGCAACTATTCTTCAATATGCTTTTGTATGATCATCTCAATCTATCACTTTCTA 3540  
Qy 4551 GGTTCCTGATTTCTTTCATAAGATTTCTTGAAGAAATGTCGAAATAATGTCGACAC 4610  
Db 3541 GGTTCCTGATTTCTTTCATAAGATTTCTTGAAGAAATGTCGAAATAATGTCGACAC 3600  
Qy 4611 TATAATCCATGGATAGCTTTTCAAGTCTTTTATGCTTTTACCTGCACTGTTGTA 4670  
Db 3601 TATAATCCATGGATAGCTTTTCAAGTCTTTTATGCTTTTACCTGCACTGTTGTA 3660  
Qy 4671 TGGATTTTCTCTTCAATATGATGAGAAATATGAGGAGCAGATCAATAAGAAAGAT 4730  
Db 3661 TGGATTTTCTCTTCAATATGATGAGAAATATGAGGAGCAGATCAATAAGAAAGAT 3720  
Qy 4731 CCCTTTTTCAGAAACCTTTTCAACGAGCTTAAATAAGGAGCTTCCAGAACCCACGAC 4790  
Db 3721 CCCTTTTTCAGAAACCTTTTCAACGAGCTTAAATAAGGAGCTTCCAGAACCCACGAC 3780  
Qy 4791 AATGAGGATGAAGATGAAGATGTCAAGAGCTTAAAGGCTCAAGAGCTGATGGGT 4850  
Db 3781 AATGAGGATGAAGATGAAGATGTCAAGAGCTTAAAGGCTCAAGAGCTGATGGGT 3840  
Qy 4851 TGGCAGTTGTGAGAGAAACCAATCATTTATGCTCAGCAATTTGCAATAAGAAATGAT 4910  
Db 3841 TGGCAGTTGTGAGAGAAACCAATCATTTATGCTCAGCAATTTGCAATAAGAAATGAT 3900  
Qy 4911 GACAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGGCACTTAATCACTCTCT 4970  
Db 3901 GACAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGGCACTTAATCACTCTCT 3960  
Qy 4971 TTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGTCCAAATGCTGTCGCAAAAGC 5030  
Db 3961 TTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGTCCAAATGCTGTCGCAAAAGC 4020  
Qy 5031 ACAATTTAATATTCTGTTGGTATATTGAACCAATTTGAGGAGGATTTTATTTAGGA 5090  
Db 4021 ACAATTTAATATTCTGTTGGTATATTGAACCAATTTGAGGAGGATTTTATTTAGGA 4080

Qy 5091 GATTATTTCTTTCAGAGACAAGTGAAGATGATGATTTCACTGAAGTGTATGGTACTGTCTCT 5150  
Db 4081 GATTATTTCTTTCAGAGACAAGTGAAGATGATGATTTCACTGAAGTGTATGGTACTGTCTCT 4140  
Qy 5151 CAGATAAAACCTTTTGGCCAGATACATCTGACAGAAATTTTGAATTTATGAGCT 5210  
Db 4141 CAGATAAAACCTTTTGGCCAGATACATCTGACAGAAATTTTGAATTTATGAGCT 4200  
Qy 5211 GTCAAGAGGATGAGTCAAGTGCATGAAAGAGTCAATAGTCAATTAACACATGCACTT 5270  
Db 4201 GTCAAGAGGATGAGTCAAGTGCATGAAAGAGTCAATAGTCAATTAACACATGCACTT 4260  
Qy 5271 GATTTTAAAGAACATCTTTCAGAGACTGTAAAGAAATCTACTTCAGAGAAATCAACGAAAG 5330  
Db 4261 GATTTTAAAGAACATCTTTCAGAGACTGTAAAGAAATCTACTTCAGAGAAATCAACGAAAG 4320  
Qy 5331 TTGTGTTTGTCTTAAGTATGCTAGGAAATCTCAGATTTACTTTGCTAGATGAACATCT 5390  
Db 4321 TTGTGTTTGTCTTAAGTATGCTAGGAAATCTCAGATTTACTTTGCTAGATGAACATCT 4380  
Qy 5391 ACAGGTATGATGCCCAAGCCAAACAGACATGTCGAGCAATTCGAACCTGCAATTTAAA 5450  
Db 4381 ACAGGTATGATGCCCAAGCCAAACAGACATGTCGAGCAATTCGAACCTGCAATTTAAA 4440  
Qy 5451 AACAGAAAGCGGCTCTTATTTCTGACCCTCACTATATGAGGAGGAGGAGGCTGTCTGT 5510  
Db 4441 AACAGAAAGCGGCTCTTATTTCTGACCCTCACTATATGAGGAGGAGGAGGCTGTCTGT 4500  
Qy 5511 GATCGAGTATGATGATGCTGCTGGCAGTTAAGATGATTCGGAACAGTACCAATCTA 5570  
Db 4501 GATCGAGTATGATGATGCTGCTGGCAGTTAAGATGATTCGGAACAGTACCAATCTA 4560  
Qy 5571 AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTTGAAGCTGATAGAAAC 5630  
Db 4561 AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTTGAAGCTGATAGAAAC 4620  
Qy 5631 CTAGAAGTATAGCCGCTTCAAGAGAAATTCAGTATATTTCCCAATTCGAAGCCGTCAG 5690  
Db 4621 CTAGAAGTATAGCCGCTTCAAGAGAAATTCAGTATATTTCCCAATTCGAAGCCGTCAG 4680  
Qy 5691 GAAAGTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTCAAGTCCCTTTCA 5750  
Db 4681 GAAAGTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTCAAGTCCCTTTCA 4740  
Qy 5751 CAATCTTTTTHAAGCTGGAAGAGCT 5777  
Db 4741 CAATCTTTTTHAAGCTGGAAGAGCT 4767

## RESULT 11

AF491842 5347 bp mRNA linear ROD 25-FEB-2003  
LOCUS Mus musculus ATP-binding cassette transporter sub-family A member 5  
DEFINITION (Abca5) mRNA, complete cds.  
ACCESSION AF491842  
VERSION AF491842.1 GI:22087247  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 5347)  
AUTHORS Annilo, T., Chen, Z.-Q., Shulenin, S. and Dean, M.  
TITLE Evolutionary analysis of a cluster of ATP-binding cassette (ABC) genes  
JOURNAL Mamm. Genome 14 (1), 7-20 (2003)  
MEDLINE 22419899  
PUBMED 12532264  
REFERENCE 2 (bases 1 to 5347)  
AUTHORS Annilo, T., Chen, Z.-Q., Shulenin, S. and Dean, M.  
TITLE Direct Submission  
JOURNAL Submitted (12-MAR-2002) Laboratory of Genomic Diversity, National

FEATURES	Cancer Institute, Frederick, MD 21702, USA	
	Location/Qualifiers	
source	1. .5347	
	/organism="Mus musculus"	
gene	/mol_type="mRNA"	
	/strain="BALB/c"	
CDS	/db_xref="taxon:10090"	
	1. _5347	
	/gene="Abca5"	
	210. .5138	
	/gene="Abca5"	
	/codon_start=1	
	/product="ATP-binding cassette transporter sub-family A member 5"	
	/protein_id="AAM90895.1"	
	/db_xref="GI:22087248"	
	/translation="MATAIRDVGVWRQTRLLLNKLYIKRTRKSSVQIPLPLPLF	
	WLLVSMHMKYEVSDIELSPKDFSLNVLGYTVTNITSSIMQVSDHLPK	
	VIIVTEYANEKELVAASLSKSNFVGVKDTMSYELRPFPMIPVSSIYMSRSGCS	
	KTCDAQYNSLFTVLQASIDAIIQLETNVSVSELSSTKAVIMGEAAVVIDTPPR	
	GVILLIVLAFSPGFLAIHIVAEKKLKEPLKIMGLDHDTAFWSWLLVASLIFL	
	MSLMAVIATASSLFFQSSIVLFLVFLIGSSVFALMLPLFKSKRHGVVREFFV	
	TVVFGVGLLIVLIESFPRSLVWFLSPLOCAFLIGIAQVMHLEDNEGALFSLNTEG	
	PYPLIITIIIMALDSVYVLLAVYLDQVIPGEFLRRSLYFLKPSYWNKRNRYEL	
	SEGININISLNEIVRPSSEFIKEAIRISGIQSKYRKKTENVEALRNLSFDIYEGQ	
	ITALLGHSGTGKSLANILGCPDPSDGPASIVGHRVSEIDEIFEARKMIGICPSODI	
	NFDVLIVENLSLAKIGIPANNIIQEVQVQLLLDMDQAIKDNOAKKLSGGQKLS	
	VGLAVLGNKILLIDPTAGMDPCSRHIVWNLKLYRKARVTVFSTHMEADILADR	
	KAVISQGMKVCDDSIPLSKMGIYRLSMYIDRYCATESLSLVRQHIIPAAALQON	
	DOQVLPLPKMDKFSGLDIIHNSLGIYSYGVSMITLEDVFLKLEVEAEIDQAD	
	YSVFTQOPREGTDSKSFDEMEQSLILSETKASSVSTMPKQVQVSTIAKPHFSLN	
	RESKSVRAVLLILLIFAVQIFMVFLHSPKNAVPIKLVPDLYFLKPGDKHKYKTS	
	LLQNSTDSDINGLIEFFAHQNTVMVFNDSDVSAAPHSALNVRSEKDVFSNVP	
	NSTMVYCLPVMENAEHNIKAYTQCLKSGLLPSAYVWQAVDIPLFVFLILMLGSL	
	FAFHGLIYFFPAFLAVFVCLIAIYVPSVILFTYIASFTFKILNTEKFPWSFIYSVAL	
	ACVAITETTFLOVAVTAVPHYFCIAIPIYPLGLCLISFIKSGMKNKNTNPNW	
	DRLLVAVMPLYCLILWIFLLOHVEKIHGGRSIRKDPFRALSOAKKNTKVPPEPINE	
	DEDEDVKAERLYKVELMGCCCEKEPAIMVGNLHKEYDDKDPKLSRKTNTKVATYIS	
	FCVKGEIIGLLGNAGKSTVINTLVGVEPTSGKIFLGDYGHSSDEDESICKMGY	
	CPQTNPLPDLTQIEHFEIYGAVKMSPGDMKEVISRIITKALDLKELHOKTVKKLPAG	
	IKRKLFPALSMLGNOVTLDEPSTGMDPRAKQHMWRAITAFKNKRAALLTTHYME	
	EAEAVCDRGAIVMSGLRCIGTQVHLKSKFGKGFLEIKLKWIEENLEIDRLQREIQY	
	IFPNASROESFSSILAPKIPKEDVQSLSQSFALKEAEAKRTFAIEEYSPSQATLEQVVF	
	ELIYKEBEEDNSCGTLASTLWERTQEDRVF"	
	BASE COUNT 1503 a 1064 c 1179 g 1601 t	
ORIGIN		
Query Match 59.6%; Score 3887; DB 10; Length 5347;		
Best Local Similarity 85.0%; Pred. No. 0;		
Matches 4381; Conservative 0; Mismatches 751; Indels 24; Gaps 2;		
Qy	991 AATAGGTTTATTTCAGAAAAATGTCCACTGCAATTTAGGAGGTTAGGAGTTCAGAGACAGA 1050	
Db	190 AGTAACTAGCTAGAAAAATCATGCTACTGCAATTTAGGATGTGGAGTTCAGAGACAGA 249	
Qy	1051 CCAGAACACTTCTACTGAGAAATTTACTTAATTAATGCGAACCCAAAGAGTAGTGTTC 1110	
Db	250 CCAGAACACTTCTACTGAAAAATTTACCTTAATTAATGCGAGCTTAAAAAAGTAGTGTTC 309	
Qy	1111 AGGAAATCTTTTCCCACTATTTTTTTTATTTGGTTAATTAATTAATGATGATGATGCATC 1170	
Db	310 AGGAAATCTTTTCCCTCTATTTTCTTATTTTGGCTGATATTAATTAATGATGATGATGCATC 369	
Qy	1171 CAAATAAGAAATATGAAGAAAGTGCTTAATATAGAACTCAATCTCTATGAGCAAGTTTACTC 1230	
Db	370 CAAATAAGAAATATGAAGAGGTATCTGATATAGAGCTCAGCCCTATGAGCAAAATTCAGCC 429	
Qy	1231 TTTCTAATCTAATTTCTTTGGATATCTCCAGTGACTAATATTAACAGCAGCATCATGACAGA 1290	
Db	430 TTTCCAAACGTTATTTCTTTGGATACATCCCGTGACTAACATTACAAGCAGCATTTATGACA 489	
Qy	1291 AAGTGCTCTGATCATCTACCTGATGTCATATTAATTAATTAATTAATTAATTAATTAATTAAT 1350	
Db	490 GGGTTTCTACCGATCATCTTCCCAAGGTTATAGTTTACTGAGAAATATCGCAAAATGAGAAAG 549	
Qy	1351 AAAATGTTAAATCATTCAGTCTCTCTAAGCCGAGCAACTTTTGTAGGTGTGTGTTTCAAAGACT 1410	
Db	550 AACTGGTAGCCGCGAAGTCTTTCTAAGTCCAGCACTTCGTAGGTGTGTGTTTCAAAGACA 609	
Qy	1411 CCATGCTCTATGAACCTTCGTTTTTTTCTGTATGATGATTCACAGTATCTTCTATATATATGG 1470	
Db	610 CCAATGCTCTATGAACCTTCGTTTTTTTCTGTAAATGAATTCAGTGTCTTCTATATATATGA 669	
Qy	1471 ATTCAAGAGCTGGCTGTTCAAAAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTCA 1530	
Db	670 ATTCAAAGAGAAAGGCTGTTCAAAGACATGTGATGCTGCTCAGTACTGTCTTTGGGGTTTA 729	
Qy	1531 CAGTTTTTAAAGACATCCATAGATGCTGCATTTATACAGTTCGAAGACCAATGTTTCTCTTT 1590	
Db	730 CAGTTCTCGAGCATCGATAGATGCTGCGCATTTATACCTCGAGACCAATGTTTCTGTGT 789	
Qy	1591 GGAAGAGCTGGAGTCAACTAAAGCTGTTTATTTATGGGAGAAACTGCTGTGTGTAGAAATAG 1650	
Db	790 GGAGCGAGCTGGAGTTCGACCAAGCTGTGATCATGGGAGAGCCGCTGTGTGGAGATTG 849	
Qy	1651 ATACCTTTCCCGAGGAGTAATTTTAAATACCTAGTTATAGCAATTTTCACTTTTGGAT 1710	
Db	850 ACACCTTCCCGAGGAGGCTCATCTCTCATCTACCTGCTCATAGCCCTTCTCGCCCTTCGGCT 909	
Qy	1711 ACTTTTGGCAATTCATATCGTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGA 1770	
Db	910 ACTTCTCGCAATCCACATCGTGGCAGAAAAAGAAAGAAAGTTTAAAGAAATTTTAAAGA 969	
Qy	1771 TAATGGGACTTCATGATACTGCTTTTGGCTTCTCTGGGTTCTTCTATATACAAGTTTAA 1830	
Db	970 TAATGGGACTTCATGACACTGCTTTTGGCTTCTCTGGGTTCTTCTGTACGCAAGCTTGA 1029	
Qy	1831 TTTTCTTTATGTCCTTCTTATGGCAGTCATTTGGCAGAGCTTCTTGTGTTATTTCTCTCAA 1890	
Db	1030 TTTTCTTTATGTCCTTCTTATGGCAGTCATTTGGCAGAGCTTCTTGTGTTATTTCTCTCAA 1089	
Qy	1891 GTAGCAGATTTGTGATATTTCTGCTTTTCTTATTTTCTTATTTGGAATATCATCTGTATTTT 1950	
Db	1090 GTAGCAGATTTGTGATATTTCTTCTTCTTATTTGGAATATCATCTGTGTTTCTGTTT 1149	
Qy	1951 CTTTAAATGTCGACACTCTTTTAAAAAATCAAAACATGTGGGAAATAGTTGAATTTT 2010	
Db	1150 CTTTAAATGTCGACACTCTTTTAAAAAATCAAAACATGTGGGAGTTCGTGAGTTT 1209	
Qy	2011 TTAATGTCGCTTTGGATTTATTTGGCTTATGATTAATCTCTATAGAAAGTTTTCCTCAAT 2070	
Db	1210 TCACCGTGTGTGTTTGGATTTTGTGGCTGCTGATTTGCTCATAGAAAGTTTTCCTCAGGT 1269	
Qy	2071 CGTTAGTGTGCTTTTCTGCT 2130	
Db	1270 CGCTGTGTGCTTTCTGCT 1329	
Qy	2131 TCATGCAATTTAGAAATTTAATGAAGGTGCTTCAATTTTCAAAATTTGACTGAGGCCCAT 2190	
Db	1330 TCATGCAATTTAGAAATTTAATGAAGGTGCTTCAATTTTCTAAATTTGACTGAAGGTCCCT 1389	
Qy	2191 ATCTCTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2250	
Db	1390 ATCTCTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1449	
Qy	2251 CTTCTATCTCTGATCAAGTCATTTCCAGGGAAATTTGGCTTACGGAGATCATCTTTATAT 2310	
Db	1450 CTTCTATCTCTGATCAAGTCATTTCCAGGGAAATTTGGCTTACGGAGATCATCTTTATAT 1509	
Qy	2311 TTTCTGAAGCTTCTATATTTGGTCAAAAGAGTAAAGAAATTTATGAGGAGTTTATCAGAGGCA 2370	
Db	1510 TTTTGAAGCCATCTGATTTGGTCAAAAGAGTAAAGAAATTTATGAGGAGTTTATCAGAGGCA 1569	
Qy	2371 ATGTTTAAATGGAATATATAGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTAT 2430	

Db 1570 ACATTAATGCAATATTAGTCTCAATGAAATGTTGAGCCGTTTCTTTCAGAAATTTATAG 1629  
Qy 2431 GAAAGAAGCATAAAGAAATAGTGGTATTTCAGAAAGACATACAGAAAGAGCGTGAATAAG 2490  
Db 1630 GGAAGNAGCTATAGAAATAGTGGTATTTCAGAAATCTATAGAAAGAAACTGAGAACG 1689  
Qy 2491 TGAAGGCTTTGAGAAATTTGCTCAATTTGACATATATAGAGGCTCAGATTACTGCTTACTTG 2550  
Db 1690 TGAAGGCTTTGAGAAATTTGCTCAATTTGACATATATAGAGGCTCAGATTACTGCTGCTGG 1749  
Qy 2551 GCCACAGTGAACAGGAAGAGTACATTGATGATATTTCTTGTGACTCTGCCCACTT 2610  
Db 1750 GCCACAGTGAACAGGAAGAGACACTGATGAATATTTCTGTGGACCGTGTCCACCT 1809  
Qy 2611 CTGATGGGTTTGGCATCTATATATGGAACAGAGTCTCAGAAATAGATGAATGTTTGAAG 2670  
Db 1810 CTGATGGGTTTGGCATCTATATATGGAACAGAGTCTCTGAATAGATGAATATTTGAAG 1869  
Qy 2671 CAAGAAATATGATGGCAATTTGTCACAGATTAGATATACACTTTGATGTTTTCACAGTAG 2730  
Db 1870 CAAGAAATATGATGGCATATGTCGCGAGTCAGATATAAACTTTGATGTTTCTGACAGTAG 1929  
Qy 2731 AAGAAATATGATGGCAATTTGTCACAGATTAGATATACACTTTGATGTTTTCACAGTAG 2790  
Db 1930 AAGAAATATGATGGCAATTTGTCGATGTCGATGTCAGATGACAGCCATCAAGATATCAAGCGNAAA 2049  
Qy 2791 AAGTGAAGAGGTTTACTAGATTTAGACATGACAGACTCAAAAGATAACCAAGCTAAAA 2850  
Db 1990 AAGTGAAGAGGTTTCTGATGTCGATGTCGATGTCAGATGACAGCCATCAAGATATCAAGCGNAAA 2049  
Qy 2851 AATTAAGTGTGFTCAAAAAGAAAGCTGTCAATTTAGGAATGCTGTTTGGGAAGCCAA 2910  
Db 2050 AGTTAAGCGTGTGACAGAAAGAAAGCTGTCTGTAGGAATGTCGAGTTCTCGGAAATCCAA 2109  
Qy 2911 AGTACTGCTGTAGATGAACCAACAGCTGGAATGAGCCCTGTTCTCGACATATTGTAT 2970  
Db 2110 AGTACTGCTGTAGACAGCTTACAGCAGAAATGGAACCCCTGCTCTCGCATATTGTTT 2169  
Qy 2971 GGAATCTTTTAAATAACAGAAAGCCCAATCGGTGACAGTGTTCAGTACTCATTTCAATGG 3030  
Db 2170 GGAATCTTTTAAAGTATAGAAAGCTAACAGATGACCGTGTGTAGTACTCACTTCATGG 2229  
Qy 3031 ATGAAGCTGACATTTTTCGAGATAGGAAGCTGTGATATACAAAGGAATGCTGAATGTG 3090  
Db 2230 ATGAGGCTGACATTTTTCGCGACAGGAAGCTGTCTATATACAAAGGAATGCTGAAGTGTG 2289  
Qy 3091 TTGTTCTTCAATGTTCTCAAAAGTAAATGGGGATCGGCTACCGCTCAGCATGTACA 3150  
Db 2290 TTGTTCTTCAATTTTCTTAAAGATTAATGGGGAATCGGCTACCGCTCAGCATGTATA 2349  
Qy 3151 TAGACAAATATTGTCACAGAAATCTCTTTCTTCACTGGTTTAAACACATATACCTGGAG 3210  
Db 2350 TAGACAGATATTGTCACAGAGTCACTGTGTCGCTGGTTTAGCAGCACATACCGCAG 2409  
Qy 3211 CTACTTTTATTAACAGAAATGACCAACACTTGTGTATAGCTTGCCTTTCAAGGACATGG 3270  
Db 2410 CCGCGCTACTGCAGCAATGACCCAGCAGCTGCTGTGTACAGCCCTGCCCTTCAAGACATGG 2469  
Qy 3271 ACAAATTTTCAGGTTTGTGTTTCTGCGCTAGACAGTCAATCAAAATTTGGGTGTCAATTC 3330  
Db 2470 ACAAATTTTCAGGCTTGTGTTTCTGCTCTAGACATTCATTCAAAACCTGGGTGTATTTCT 2529  
Qy 3331 ATGGTGTTCATGACCACTTTTGAAGACGTTATTTTAAAGCTAGAAATTTGAAGAGAAA 3390  
Db 2530 ATGGTGTTCATGACCACTTTGAAGATGTATTTCTTAAAGCTAGAAATTTGAAGAGAAA 2589  
Qy 3391 TTGACCAAGCAGATTTAGTGTATTTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAA 3450  
Db 2590 TTGACCAAGCAGACTAGTGTATTTTACACAGCAGCCGCGGAGGAGGAACAGATTCAA 2649  
Qy 3451 AATCTTTTGAATGAAGAACAGAGTCTTATTTCTGTAATTTCTGTAACCAAGGCTTCTTAG 3510  
Db 2650 AATCTTTTGAATGAAGAACAGAGTCTTATTTCTCTGTAATTTCTGTAACCAAGGCTTCTTAG 2709

Qy 3511 TGAGCACCATGAGCCCTTTGGAAACCAACAGATGTATACAATAGCAAAAGTTTCATTTCTTTA 3570  
Db 2710 TGAGCACCATGAGCCCTTGGAAGCAAGTGTCTAGGATGCAAAAGTTTCATTTCTCTCT 2769  
Qy 3571 CTTTGAACCGTGAAGTAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCAC 3630  
Db 2770 CATTTGAACCGAAGAAACAAATCAGTGGCGGTGTGTGCTCTCTTTTAAATTTTTCG 2829  
Qy 3631 CAGTTTCAGATTTTTATGTTTTGTTTTCACACTCTTTTAAATAATGCTGTGTTTCCCATCA 3690  
Db 2830 CAGTTTCAGATTTTTATGTTTTTCTCCATCATCTTTTAAATAATGCTGTGTTTCCCATCA 2889  
Qy 3691 AACTTTTCCAGACTTATATTTTCTAAAAACCTGGAGACAAACCAATAAATACAAAACAA 3750  
Db 2890 AACTGTTTCCAGACTTGTATTTCTTAAAGCTGGAGATAAACCTCATATAACAAAACAA 2949  
Qy 3751 GTCTGCTCTTCAAAATCTGCTGACTCAGATATCAGTATCATTTATAGCTTTTTCACAA 3810  
Db 2950 GCTGCTGCTTCAAAATCTTACTGACTCAGATATCAATGGTCTTATTTAGTGTTTTTCGAC 3009  
Qy 3811 GCCAGAACATAATGGTGACGATTAATGACAGTCACTATGATCGGTGCTCCCATCA 3870  
Db 3010 ACCAGAACATAATGGTGGCAATGTTTAAATGACAGTCACTATGCTGCTGCTCTCACA 3069  
Qy 3871 GTGCGGCTTTAAATGTGATGCAATTCAGAAAAAGGACTATGTTTTTGAGCTGTTTTCAACA 3930  
Db 3070 GTGCGGCTCTCAATGTGGTGGTCTCTGAAAAAGGACTATGTTTTTCTGCTGTTTTCAACA 3129  
Qy 3931 GTACTATGTTTATTTCTTTACCTATATTAGTGAATATCATTTAGTAATCTACTATCTTTATC 3990  
Db 3130 GTACTATGTTTATTTGTTTCCAGTCAATGATGAACATCATTTAGTAATCTACTATCTTTATC 3189  
Qy 3991 ATTTAAATGTGACTGAACCAATCCAGATCTGGAGTACCCCATTTCTTTCAAGAAATTACTG 4050  
Db 3190 ATTTAAATGTGACTGAAGCCATCCAGCTGGAGTACCCCGTTTCATTTCAAGAAATTACTG 3249  
Qy 4051 ATATAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTTGGTTGGAATCATTTGTTACTG 4110  
Db 3250 ACATTTGTTTTTAAATTTGAGCTATATTTTCAAGCAGCTTTTGGTTGGAATCATTTGTTACTG 3309  
Qy 4111 CAATGCCACTTACTTTTGGCATGGAATAATGCAGAGAAATCATTAAGATCAAAAGCTTATCTC 4170  
Db 3310 CAATGCCACTTACTTTTGGCATGGAATAATGCAGAGAAATCATTAAGATCAAAAGCTTATCTC 3369  
Qy 4171 AACTTAAACTTTCAGGCTTTTCCCATCTGCATATTTGGATGGAAGTGGACAGCTGTGTTGATA 4230  
Db 3370 AACTTAAACTTTTGGGCTCTTTTGGCCATCAGCCCTATTTGGGTGGACAGCTGTGTTGATA 3429  
Qy 4231 TCCCTTATTTTTTATCATTTTATTTGATGCTAGGAAGCTTACTTGGCAATTTTCATTTATG 4290  
Db 3430 TTCCCTTGTTTTTTGTGTTCTCATTTTGTGATGCTGGAGAGTTTATTTGCAATTTTCATCATG 3489  
Qy 4291 GATTATTTTTTATCTGTAAGTTCCTTGTGCTGTTTTTGGCTTTATTTGTTATGTTTC 4350  
Db 3490 GACTGTATTTTTTATCCTGCAAAATTTCTTGTGCTGTTTTTGGCTCAITTTGCTTATGTCG 3549  
Qy 4351 CATCAGTATTTCTGTTTCACTTATTTGCTTCTTCACTTTTAAAGAAATTTTAAATACCA 4410  
Db 3550 CCTCGCTCATTTCTGTTTCACTGATAGCTTTCGTTCTCTTCAAGAAAAATTTTAAATACCA 3609  
Qy 4411 AAGAAATTTTGGTTCATTTATCTATTCTGTGCGAGCGTTGCTTGTATTTGCAATCACTGAAA 4470  
Db 3610 AGGAATTTTGGTTCATTTATCTATTCTGTGACCGCATTTGGCTGTGTGGCAATCAACGAAA 3669  
Qy 4471 TAACTTTTATTTGGATACAAATTTGCAACTATTTCTTCAATGCTTTTGTATCATCA 4530  
Db 3670 CAACTTTCTTTCTGCAATATGCAAGTTTACGCTGCTCTTTTCAATACACTTCTGCTATGCA 3729  
Qy 4531 TTCCAATCTTATCCACTTCTAGTTGCTGATTTCTTTTCAATAAGATTTCTTTGGAAGATG 4590  
Db 3730 TTCCAATCTACCCCTCTCTGGGTGTGTGATTTCTTTTCAATAAGGGTCTTTGGAAGATA 3789





Db 1636 CTGATGGTTTCTTCTATATATGGACACAGAGTCTCTGAAATAGATGAATGTTTGAAG 1695  
Qy 2671 CAAGAAAATGATTTGGCAATTTGTCACAGTTAGATATACACTTTGATGTTTGGACAGTAG 2730  
Db 1696 CGAGAAAATGATTTGGCATATGTCACAGTCAGATGAACCTTTGATGTTTCTGACAGTAG 1755  
Qy 2731 AAGAAAATTTATCAATTTTCGCTTCAATCAAGGGATACAGCCACAAATATATATCAAG 2790  
Db 1756 AAGAAAATTTATCAATTTTGGCTTCAAGTCAAGGAATACAGCCACAAATATATTTCAAG 1815  
Qy 2791 AAGTCAGAGAGTTTACTAGATTTAGACATGACAGACTATCAAGATATAACAAAGCTAAAA 2850  
Db 1816 AAGTCAGAGAGTTTACTGATTTGACATGACAGCCATCAAGATATAACAAAGCGAAA 1875  
Qy 2851 AATTAAGTGTGTGTCAAAAGAAAGCTGTCAATAGAAATGCTGTTTGGGAAACCCAA 2910  
Db 1876 AACTAAGTGTGTGTCAAAAGAAAGCTGTCTTTAGGAATGTCAGATTTCTTTGGGAATCCAA 1935  
Qy 2911 AGATACCTGCTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATTTGTAT 2970  
Db 1936 AGATCCTCCTGCTAGATGAGCCGACAGCAAGATGGACCCCTGCTCTCGCCACATTTGTTT 1995  
Qy 2971 GGAATCTTTTAAAATACAGAAAAGCCCAATCGGCTGACAGTGTTCAGTACTCATTTTCATGG 3030  
Db 1996 GGAATCTTCTCAAGTACAGGAGGCTTAACGAGTGAAGTCTTCAGCACTCACTTCATGG 2055  
Qy 3031 ATGAAGCTGACATTTCTTGGATAGAGAAAGCTGTGATATACAAAGGAATGCTGAAATGTG 3090  
Db 2056 ATGAGGCCGACATTTCTTGGCCGACAGGAAAGCTGTCTATATACAAAGGAATGCTGAAATGTG 2115  
Qy 3091 TTGGTTCTTCAATGTTCTCTCAAGATTAATGGGGATCGCTACCGCTGACAGATGTACA 3150  
Db 2116 TTGGTTCTTCAATTTTCTGAAAAGTAAATGGGGAATCGGCTACCGTCTGAGCATGTATA 2175  
Qy 3151 TAGACAAATATTTGGCCACAGAAATCTCTTTCTTCACTGGTTTAAACACATATACCTGGAG 3210  
Db 2176 TAGACAGTACTGTGCCACAGAGTCGCTCTCGTCGCTGGTTAGACAGCACATCCCTGGG 2235  
Qy 3211 CTACTTTATTAACAGAAATGACCAACAACTTGTGTATAGCTTGGCTTCAAGGACATGG 3270  
Db 2236 CCGCTCTGCTGCAGCAATGACCCAGCAGATTTGTCTATAGCTTGGCCCTTCAAGACATGG 2295  
Qy 3271 ACAAATTTTCAAGTTTCTGCTTCTGCTTACAGTCACTTCAAAATTTGGGTCATTTCTTT 3330  
Db 2296 ACAAATTTTCAAGCTTGTCTCTGCTCTAGACATTCATCAAACTTGGGTCATTTCTTT 2355  
Qy 3331 ATGTTGTTTCCATGACGACATTTGGAAGACGTTATTTTAAAGCTAGAAATGGAAGCAGAA 3390  
Db 2356 ATGTTGTTTCCATGACGACATTTGGAAGACGTTTAAAGCTAGAAATGGAAGCAGAA 2415  
Qy 3391 TTGACCAAGCAGATTAATAGTATTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAA 3450  
Db 2416 TTGACCAAGCAGATTAATAGTATTTACTCCAGCAGCCACTGGAGGAAGAAACAGATTCAA 2475  
Qy 3451 AATCTTTTGTATGAATGGAAAGCAGCTTACTTATTTCTTCTGAAACCAAGCTTCTCTAG 3510  
Db 2476 AGTCTTTTGTATGAATGGAAAGCAGATTTACTTATTTCTCTGAAACCAAGCTTCTCTAG 2535  
Qy 3511 TGAGCACCATGAGCTTTTGGAAAACAAAGATGTATACAAATAGCAAAAGTTTCATTTCTTTA 3570  
Db 2536 TGAGCACCATGAGTCTCTGGNAGCAGCAAGTATCTACAAATAGCAAAAGTTTCATTTCTCT 2595  
Qy 3571 CTTTGAACCGTGAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTCTTCA 3630  
Db 2596 CACTTGAAGCGAGAAAGCAAAATCAGTGGCTCTGTGTGCTTCTGCTTTTAAATTTTCTTGG 2655  
Qy 3631 CAGTTTCAGATTTTATGTTTGTGTTTCTCATCTCTTTTAAATGCTGTGGTCCCATCA 3690  
Db 2656 CAGTTTCAGATTTTATGTTTGTGTTTCTCATCTCTTTTAAATGCTGTGGTCCCATCA 2715  
Qy 3691 AACTTGTTCAGACTTATATTTTCTTAAACCTGGAGACAAACCATATAAAATACAAACAA 3750

Db 2716 AGCTGTTTCCAGACTTGTATTTTCTTAAACCTGGAGATAAAACCTCATATAAACAACAA 2775  
Qy 3751 GTCTGCTTCTTCAAAATTTCTGCTGACTCAGATATCAGTGTATCTTATAGCTTTTTCACAA 3810  
Db 2776 GCTGCTGCTTCAAAATTTCTGATTCAGATATCAATGATCTTATGACTTTTTCACAC 2835  
Qy 3811 GCCAGAAACATATATGCTGACGATGATTAATGACAGTGACTATGATCCGCTGGCTCCCATTA 3870  
Db 2836 AGCAGAACATATTTGCGCATGTTTAAATGACAGTGACTATGCTGCTGCTCTCATTA 2895  
Qy 3871 GTGCGGCTTAAATGTGATGCAATTCAGAAAAGGACTATGTTTTTGGAGCTGTTTTCACAA 3930  
Db 2896 GTGCGGCTCTCAATGTGGTGCAGTCAGAAAAGGACTATGTTTTTACTGCTGTTTTTCAACA 2955  
Qy 3931 GTACTATGTTTATTTCTTTTACCTATATAGTGAATATCATATAGTAACTACTATCTTATC 3990  
Db 2956 GTACTATGTTTATTTCTTTTGGCAGTAATGATGAACATCATATAGTAACTACTATCTATC 3015  
Qy 3991 ATTTAAATGTGACTGAAACCATCCAGATCTGAGTACCCCATTTCTTCAAGAAATTTACTG 4050  
Db 3016 ATTTAAATGTGACTGACACTATCCAGATCTGAGTACCCGTTTCAATTCAGGAAATTTACTG 3075  
Qy 4051 ATATAGTTTTTAAATTTGAGCTGTATTTTCAAGAGCTTTGCTTGGAAATCATTTGTTACTG 4110  
Db 3076 ACATTTGTTTTTAAAGTTGAGCTATATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTG 3135  
Qy 4111 CAATGCCACTTACTTTTGCATGGAATGCAAGATCATAAAGCTTATACTTCAAGCTTATACTC 4170  
Db 3136 CAATGCCACTTACTTTTGCATGGAATGCAAGATCATAAAGCTTAAAGCTTATACTC 3195  
Qy 4171 AACTTAAACCTTTCAAGTCTTTTGCATATTTGGAATGCAAGCTTGTGTTGATA 4230  
Db 3196 AGCTTAAACCTTTCAAGTCTTTTGCATCGCCTACTGGATGCAAGCTGTTGTTGACA 3255  
Qy 4231 TCCCTTATTTTATTTTATCTTTATTTGATGCTAGGAGCTTACTGGCAATTTTCAATG 4290  
Db 3256 TTTCCCTGTTTTTGTGTTCTGACTTTTGAATGCTGGAAGTTTATTTGCTTTTCACTCAG 3315  
Qy 4291 GATTAATTTTTTATCTGTAAGTTTCTTGTGCTGTTTTTTTGGCTTATTTGTTGTTGTC 4350  
Db 3316 GACTGTATTTTTTATCTGTAAAGTTTCTTGTGCTGTTTTTGGCTTCTGCTTATGTCG 3375  
Qy 4351 CATCAGTTATTTCTGTTTCACTTATTTGCTTTTCACTTAAAGAAATTTTAAATATCCA 4410  
Db 3376 CCTCGTTATTTCTGTTTCACTTATATAGCTTCTGTTCACTTAAAGAAATTTTAAATATCCA 3435  
Qy 4411 AAGAAATTTTGTGCTATTTATCTGTTGAGCTTGTGCTGCTGCTTGTGCTGCTGCTGCTG 4470  
Db 3436 AGGAAATTTTGTGCTATTTATCTTGTGACAGCATTTGGCTTGTGCTGCTGCTGCTGCTG 3495  
Qy 4471 TAACTTCTTTATGGATACACAAATTTGCAACTTATTTCTTCAATATGCTTTTCTGATATCA 4530  
Db 3496 TAACTTCTTCTGCTGCTATGGATTTACGGCTGTTTTTCAATACACTTCTGCTGCTGCTGCTG 3555  
Qy 4531 TTTCCAATCTATCCACTTCTAGCTTCTGCTGCTTCTTCAATAAGAAATTTCTTGGAAAGATG 4590  
Db 3556 TCCCAATCTACCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3615  
Qy 4591 TACGAAAAATGTGGACACTATTAATCCATGGGATAGGCTTTTCACTGCTGCTGCTGCTGCTG 4650  
Db 3616 TACCAAAAACTGAGAAATGCTCAATCCCTGGGACAGACTTTTGTGCTGCTGCTGCTGCTGCTG 3675  
Qy 4651 CTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4710  
Db 3676 CTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3735  
Qy 4711 GCAGATCAATAAGAAAAAGATCCCTTTTTCAGAAAACTTTTCAACGAGTCTTAAAAATATGGA 4770  
Db 3736 GCAGATCAATAAGAAAAAGATCCCTTTTTCAGGGCCCTCTTCAAAAAAGCCAAACATAAGA 3795  
Qy 4771 AGCTTCCAGAACCAACAGCAATGAGGATGAAGATGAAGCTGCAAGCTGAAAGACTAA 4830  
Db 3796 AGTTTCCAGAACCAACCTATCAATGAGGATGAAGATGAAGCTGCAAGCTGAAAGCTGGA 3855







LLQNSTDSNDINGLIBFAHONIMVAMFNDSDYVSAAPHSAAALNVVRSEKDYVFSVAVF  
NSTWVCLPMMNINISNYLHNTVEAIGTQSTWFOIETDIVDFKIELYFOAALIGL  
IVTAMPYFPAWENAEHNIKAYTOLKSLGPSAYVQVAVDIPLFVVLILMLGSL  
FAHHGLYFPAPKFLAVFCLIAVPSVILFTYIASTPKILINTKEPWSFYISVTAL  
ACVAITETTFLOAVTAVPHYTFCAIPIYPLGLCLISFPIKGSWKNMPKENTINPW  
DRLLVAVIMPYQCIILWIFLLOHYEKIHGSRIRKDPFPRALSOAKNRKKFPPELINE  
DEDEDVKAERILKVLKNGAGKSTVINTLVGDVEPTSGIKIPLGDYSHSSDEDEIKCMGY  
FCVKKGRIILGLLGNAGKSTVINTLVGDVEPTSGIKIPLGDYSHSSDEDEIKCMGY  
CPTNPLPWLDTLOEHEPEIYVAGKMSPGDMKEVIRITKALDLKSHLOKTVKKLPAK  
IKKALCFALSMLENGPOVTLDERSTGMDPRAKHMRIRATFAPKPKRAALLTHYME  
EBAVCDFALVMSVGGURCITQVOHLKSPKGFYFLEIKLKWIMENLEIDRLOREIYQ  
IFPNARQESFSSILAFKPKEDVQSLSQSPAKLEBAKRTFAIEEYSFSQATLEQVVF  
ELTKQEEDENSCTLASTLWERTQDRVVF"  
BASE COUNT 1416 a 958 c 1075 g 1480 t  
ORIGIN

Query Match 59.3%; Score 3868.8; DB 10; Length 4929;  
Best local Similarity 86.5%; Pred. No. 0;  
Matches 4266; Conservative 0; Mismatches 663; Indels 0; Gaps 0;

QY	1011	ATGTCCTCACTCAATAGGGAGTAGGAGTTTGGAGACAGACACACTTCTACTGAAG	1070
DB	1	ATGGCTACTCAATAGGGAGTAGGAGTTTGGAGACAGACACACTTCTACTGAAG	60
QY	1071	AATTACTTAATAAATGCAGAACCAAAAGAGTAGTCTTCAGGAAATCTTTTCCACTA	1130
DB	61	AATTACTTAATAAATGCAGACTAAAAGTAGTCTTCAGGAAATCTTTTCCCTCTA	120
QY	1131	TTTTTTTTTATTTGGTTTAATTAATTAATGATGATGATCCAAATAGAAATAGAGAA	1190
DB	121	TTTTTCTTATTTGGCTGATATTAGTTAGCATGATGATCCAAATAGAAATAGAGAG	180
QY	1191	GTGCTTAATAGAACTCAATCTTATGACAGAGTTTACTCTTCTTAATCTTAATCTTGA	1250
DB	181	GTATCTGATATAGAGCTCAGCCCTATGGACAAATTCAGCCCTTTCCAACTGTTTCTTGA	240
QY	1251	TATACTCCAGTGACTAATATTACAGCAGCATCATGACAGAAAGTGCTACTGATCATCTA	1310
DB	241	TACACTCCGTGACTACATTAACAGCAGCATTAATGACAGAGGTTTCTACCGATCATCTT	300
QY	1311	CCTGATGCTAATATTCTGAAGAAATATACAAATGAAAGAAATGTTTAAACATCAAGTCTC	1370
DB	301	CCCAAGGTTATAGTTACTGAAGAAATACGCAAAATGAGAAAGAACTGGTAGCCGCAAGTCTT	360
QY	1371	TCTAAGCGGACCACTTTGTAGGTGCTGTTTTCAGAGCTCCATGCTCTATGAACTTCGT	1430
DB	361	TCTAAGTCCAGCACTTCGTAGGTGCTGTTTTCAGAGCACCATGCTCTATGAACTTCGT	420
QY	1431	TTTTTCTCTGATATGATTCAGATATCTTCTATTATATGATTCAGAGCTGGCTGTCTCA	1490
DB	421	TTTTTCTCTGATATGATTCAGATATCTTCTATTATATGATTCAGAGAGGCTGTCTCA	480
QY	1491	AAATCATGTGAGGCTGCTAGTACTGCTCTCAGTTTTCACAGTTTTCACAGCATCCATA	1550
DB	481	AAGACATGTGATGCTGCTCAGTACTGCTCTTGGGTTTACAGTTCTGCGAGCATCGATA	540
QY	1551	GATGCTGCCATTATACAGTTGAAGACCAATGTTTCTCTTTGGAAGAGCTGGAGTCAACT	1610
DB	541	GATGCTGCCATTATACAGCTGAAGACCAATGTTTCTGTTGGAGCGAGCTGGAGTCGACC	600
QY	1611	AAAGCTGTTATTATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTTCCCGAGGAGTA	1670
DB	601	AAAGCTGTATATGGAGAGGCGCTGTTGTGAGAAATGACACTTCCCGAGGAGGTC	660
QY	1671	ATTTTAAATACCTTAGTATAGCAATTTTCACTTTTTCGATATCTTTTGGCAATTCATATC	1730
DB	661	ATCCTCATCTACCTCGTCATAGCCTTCTCGCCCTTCGGCTACTTCTCGGCAATCCACATC	720
QY	1731	GTACGAGAAAGAAAGAAATTAAGAAATTTTAAAGATAATGGACTTCATGATACT	1790
DB	721	GTGCGAGAAAGAAAGAAAGTTTAAAGAAATTTTAAAGATAATGGACTTCATGACACT	780
QY	1791	GCCTTTTGGCTTCTCCTGGGTTCTTCTATATACAAAGTTTAAATTTTCTTATGTCCTTCTT	1850

DB	781	GCTTTTGGCTTCTCCTGGGTTCTTCTGACGCAAGCTTGATTTTCTTATGTCCTCTT	840
QY	1851	ATGCGAGTCATTCGACAGCTTCTTTGTTATTTCTCTCAAGTAGCAGATTCGTGATATTT	1910
DB	841	ATGGCTGTCATCGCAACAGCTTCTTCGTTATTTCCCTCAGAGTAGCAGCATTTGTGATCTT	900
QY	1911	CTGCTTTTTTCTTTATAGGATTTATCATCTGTATTTTTTGTCTTTAATGCTGACACTCTT	1970
DB	901	CTACTGTTCTTCTTATATAGGATTTGTCATCTGTGTTTTTGTCTTTAATGCTGACGCTCTT	960
QY	1971	TTTTAAAAATCAAAACATGTGGGAATAGTTGAATTTTTTTGTACTGTGGCTTTTGATTT	2030
DB	961	TTTTAAAAATCAAAACACGCTGGGAGTCGTTGAGTTTTTTGTCCACGCTGTGTGTGATTT	1020
QY	2031	ATTGGCTTATGATAATCTCTCATAGAAATTTTCCAAATCGTTAGTGTGGCTTTTTCAGT	2090
DB	1021	GTGGCGCTGCTGATTTGCTCATAGAAATTTTCCCGAGCTGCTGGTGTGGCTCTTTCAGT	1080
QY	2091	CTTTTCTGTCATCTGATCTTTTGTGATTTGATTTGCACAGGTCATGCAATTTAGAAGATTTT	2150
DB	1081	CTTTTCTGTCAGTGTGCTTCTGATTTGGATTTGCACAGGTCATGCAATTTAGAAGATTTT	1140
QY	2151	AATGAAGTGTCTTCAATTTTCAAAATTTGATCTGACGCGCCATATCTCTAATTAATTAATTT	2210
DB	1141	AATGAAGTGTCTTCAATTTTCAAAATTTGATTTGATTTGATTTGATTTGATTTTACTATTT	1200
QY	2211	ATCATGCTCACAATTTAATAGTATATTTCTATGCTCTTGGCTGTCTATCTTTCATCAAGTC	2270
DB	1201	ATCATGCTCACAATTTAATAGTATATTTCTATGCTCTTGGCTGTCTATCTTTCATCAAGTC	1260
QY	2271	ATTCCAGGGAAATTTGGCTTACGAGATCATCTTTATATTTTCTGAAGCTTTCAATATGG	2330
DB	1261	ATTCCAGGGAAATTTGGCTTACGAGATCATCTTTATATTTTCTGAAGCTTTCAATATGG	1320
QY	2331	TCAAAGAGTAAAGAAATTTATGAGGATTTATGAGGGCAATGTTTAAATGGAATATTAAT	2390
DB	1321	TCAAAGAGTAAAGAAATTTATGAGGATTTATGAGGGCAATGTTTAAATGGAATATTAAT	1380
QY	2391	TTTAGTGAATTTTTCAGCCAGTTTCTTCAGAAATTTGTAGGAAAGAGCCATATAGAATTT	2450
DB	1381	CTCAATGAATTTTTCAGCCAGTTTCTTCAGAAATTTGTAGGAAAGAGCCATATAGAATTT	1440
QY	2451	AGTGGTATTTCAGAGACATACAGAAAGAGGTTGAAATTTGTGAGGCTTTTGAGAAATTTG	2510
DB	1441	AGTGGTATTTCAGAGACATACAGAAAGAGGTTGAAATTTGTGAGGCTTTTGAGAAATTTG	1500
QY	2511	TCATTTGACATATATGAGGTCAGATTTACTGCTTTTACTGTCCTGCTGGGCCACAGTGGAA	2570
DB	1501	TCATTTGACATATATGAGGTCAGATTTACTGCTTTTACTGTCCTGCTGGGCCACAGTGGAA	1560
QY	2571	AGTACATTTGATGAATATTTTCTGTCGACTCTGCCACCTTCTGATGGGTTTTCATCTATA	2630
DB	1561	AGTACATTTGATGAATATTTTCTGTCGACTCTGCCACCTTCTGATGGGTTTTCATCTATA	1620
QY	2631	TATGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAATGATTTGGCAAT	2690
DB	1621	TATGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAATGATTTGGCAAT	1680
QY	2691	TGTCACACAGTATATACATTTTGTATGTTTTGACAGTAGAAGAAATTTTCAATTTTGG	2750
DB	1681	TGTCACACAGTATATACATTTTGTATGTTTTTGTATGTTTTTGTATGTTTTTCAATTTTGG	1740
QY	2751	GCTTCAATCAAGGNTTACCAGCCCAATATATACAGAGTGCAGAGGTTTTCATCTATA	2810
DB	1741	GCTTCAATCAAGGNTTACCAGCCCAATATATATATTTCAAGAGAGTGCAGAGGTTTTCATCTATA	1800
QY	2811	GATTTAGACATGCAGATATCAAGATACCAAGCTTAAAGAAATTAAGTGGTGGTCAAAAA	2870
DB	1801	GATTTAGACATGCAGATATCAAGATACCAAGCTTAAAGAAATTAAGTGGTGGTCAAAAA	1860
QY	2871	AGAAAGCTGTCTAATAGGAATTTGCTGTTTGGGAAACCCCAAGATATCTGCTGTAGATGAA	2930

Db 1861 AGGAAGCTGCTGTAGGAAATTGCAAGTTCTCGGGAATCCAAGATACCTCTGCTAGACGAG 1920  
Qy 2931 CCAACAGCTGGAATGGACCCCTGTCTCGACATATCTATGGAATCTTTTAAAAATACAGA 2990  
Db 1921 CCTACACAGGAATGGACCCCTGCTCTCGCATATTTGTTGGAACTCTTCTAAAGTATAGA 1980  
Qy 2991 AAAGCCAAATCGGGTGAACAGTGTTCAGTACTCATTTTCATGATGAAGTGCATCTTTGCA 3050  
Db 1981 AAGGCTAAACAGAGTGACCGTGTTAGTACTCACTTCATGATGAGGCTGACATCTTTGCC 2040  
Qy 3051 GATAGGAAGCTGTGATACACAGGAATGCTGAATGCTGTGGTCTTCAATGTCTCTC 3110  
Db 2041 GACAGGAAGCTGTCAATACACAGGAATGCTGAAGTGTGTGGTCTTCAATTTTCCCTA 2100  
Qy 3111 AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGCGCAC 3170  
Db 2101 AAGAGTAAATGGGGAATCGGCTACCGCTGAGCATGTATATAGACATGACTGTGCCACA 2160  
Qy 3171 GAATCTCTTTCTCACTGGTTAAACAAATATATACCTGGAGCTACTTTTATTTACACAGAAT 3230  
Db 2161 GAGTCACTGTCGTGCTGGTTAGGCAGCACATACCCGACGCGCTACTCTGACAGCAAT 2220  
Qy 3231 GACCAACAACTTGTGTATAGCTTCCCTTCAAGGACATGACAAATTTTCAGGTTTGT 3290  
Db 2221 GACGAGAGCTCGTGTACAGCTGCTCTTCAAGACATGACAAATTTTCAGGCTTGT 2280  
Qy 3291 TCTGCCCTAGACAGTCAATCAAAATTTGGGTGTCAATTTCTTATGGTGTTCATGACGACT 3350  
Db 2281 TCTGCTCTAGACATTCATTAACCTGGGTGTATTTCTTATGGTGTTCATGACACAA 2340  
Qy 3351 TTGGAAGACGTATTTTTAAAGCTAGAAATTGAAGCAGAAATTTGACCAAGCAGATTAAGT 3410  
Db 2341 TTGGAAGATGTATTTTAAAGCTAGAAATTGAAGCAGAAATTTGACCAAGCAGATTAAGT 2400  
Qy 3411 GTATTTACTGACGACCACTGGAGGAAGAAATGGAATCAAAATCTTTTGTATGAATGGAA 3470  
Db 2401 GTATTTACACAGCAGCGCGGAGGAAGAAACAGATTTCAAAATCTTTTGTATGAATGGAA 2460  
Qy 3471 CAGAGCTTACTTATTTCTTCTGAAACCAAGCTTCTCTAGTGAGCACCATGAGCCTTTGG 3530  
Db 2461 CAGAGTTTACTTATTTCTCTGAAACCAAGCTTCTCTAGTGAGCACCATGAGCCTCTGG 2520  
Qy 3531 AAACCAACAGATGTATACAAATAGCAAAAGTTTCATTTCTTTTACCTTTGAAACGCTGAAAGTAAA 3590  
Db 2521 AAGCAGCAAGTGTCTACGATGCAAAAGTTTCATTTCTCTCTCATTTGAAACGAGAAACAA 2580  
Qy 3591 TCAGTGAGATCAGGTGTCTTCTGCTTTTAAATTTTTCACAGTTCAAGTTTATGTTTT 3650  
Db 2581 TCAGTGCGCGTGTGTGCTTCTGCTTTTAAATTTTTCAGTTTCAGATTTTATGTTTT 2640  
Qy 3651 TTGGTTTCACTCTTTTAAATGCTGTGTTCCCATCAAACTTGTTCAGACTTATAT 3710  
Db 2641 TTTCTCCATCACTTTTAAATGCTGTGTTCCCATCAAACTGTTTCAGACTTGTAT 2700  
Qy 3711 TTTCTAAACCTGGAGACAAACCATATAAATACAAAACAAAGTCTGCTTTCTTCAAAATTC 3770  
Db 2701 TTCTAAAGCCTGGAGATAAACCCTCATATAACAAACAAAGCCTGCTCTTCAAAATTC 2760  
Qy 3771 GCTGACTCAGATATCAGTGATCTTATTTAGCTTTTTCACAGCCAGAAACATAAATGGTGACG 3830  
Db 2761 ACTGACTCAGATATCAATGCTCTTATTTAGTGTGTTTTTTCACACCAAGACATAAATGGTGCA 2820  
Qy 3831 ATGATTAATCAGAGTACTATGATCGTGGCTCCCATAGTCGGCTTTAAATGTGATG 3890  
Db 2821 ATGTTTAAATGACAGTACTATGCTGCTCTCTCACAGTGGCTCTCAATGTGGTG 2880  
Qy 3891 CAATCAGAAAAGGACTATGTTTTTTCAGCTGTTTTTCAACAGTACTATGGTTTATCTTTA 3950  
Db 2881 CGTTCTGAAAAGGACTATGTTTTTTCAGCTGTTTTTCAACAGTACTATGGTTTATGTTG 2940  
Qy 3951 CCTATATTAGTGAATATCATTAGTAACTACTATCTTTATCATTTTAAATGTGACTGAAACC 4010  
Db 2941 CCACTCATGATGAACATCAATTAGTAACTACTATCTTTATCATTTTAAATGTGACTGAAGCC 3000

Qy 4011 ATCCAGATCTGGAGTACCCCAATCTTTCAAGAAATTAAGTATAGTATTTTAAAAATTGAG 4070  
Db 3001 ATCCAGACCTGGAGTACCCCGTTCATTTCAAGAAATTAAGTATAGTATTTTAAAAATTGAG 3060  
Qy 4071 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCCACTTACTTTGGCC 4130  
Db 3061 CTATATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCCACTTACTTTGGCC 3120  
Qy 4131 ATGGAATAATGCAGAGAAATCATAGATCAAAAGCTTTATCTCAACTTAAACCTTTTCAAGGCTTT 4190  
Db 3121 ATGGAATAATGCAGAGAAATCATAGATCAAAAGCTTTATCTCAACTTAAACCTTTTCAAGGCTTT 3180  
Qy 4191 TTGCCATCTGCAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTT 4250  
Db 3181 TTGCCATCAGCCTATTTGGGTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTT 3240  
Qy 4251 CTTATTTTGAATGCTAGGAAGCTTTACCTGGCAATTTTCAATTTGGAATTTTATTTTATCTGTA 4310  
Db 3241 CTGATTTTGAATGCTGGGAAGTTTATTTGCAATTTTATGCAATTTTATGCAATTTTATGCAATTT 3300  
Qy 4311 AAGTCTCTGCTGCTGTTTTTGGCTTATTTGTTTATGTTTCCATCAGTATTTCTGTTTCACT 4370  
Db 3301 AAATTTCTGCTGCTGTTTTTGCCTCATTTGCTTATGCTTCCCTCGTCAATTTCTGTTTCAAG 3360  
Qy 4371 TATATTTGCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTTATC 4430  
Db 3361 TACATAGCTTCTGTTCACTTTCAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3420  
Qy 4431 TATCTGTGGCAGCTTGTGTTGCAATCCTGAAATATCTTTCTTTTATGGGATAC 4490  
Db 3421 TATCTGTGACCGCAATTTGGCTTGTGGCAATCACGGAACAACTTTCTTTCTGCAATAT 3480  
Qy 4491 ACAATTTGCAACTATTTCTTCAATTTGTTTGTATCATCATTTCCAATCTCACTTCTA 4550  
Db 3481 GCAGTTTACCGCTGCTTTTCAATCACCTTCTGATAGCCATTTCCAATCTCACTTCTCTG 3540  
Qy 4551 GGTGCTGCTGATTTCTTTTCAATTTTAAAGAAATTTCTTGAAGAAATGTCAGAAATAATTTGTCACAC 4610  
Db 3541 GGTGCTGCTGATTTCTTTTCAATTTTAAAGGTTCTTGAAGAAATGTCAGAAATAATTTGTCACAC 3600  
Qy 4611 TATAATCTCATGGATAGGCTTTTCAAGTGTGTTTATATGCTTTTACCTGAGTGTGTAATG 4670  
Db 3601 TACAATCTCTGATAGACTTTTGTGTTGTTATCATGCTTACCTGAGTGTGTAATG 3660  
Qy 4671 TGAATTTTCTCTTCAATACTATGAAAAAATATGAGGCGAGATCAATTAAGAAAGAT 4730  
Db 3661 TGAATTTTCTCTTCAATCACTATGAAAAAATATGAGGCGAGATCTATTAAGAAAGAT 3720  
Qy 4731 CCTTTTTCAGAAACCTTTTCAACGAAAGTCTTAAATAAGGAAGCTTCCAGAACCAACAGAC 4790  
Db 3721 CCATTTTTCAGGCTTTTTCAGAAAGGCGCAAAATTAAGAAAGTTTCCAGAACCAACAGAC 3780  
Qy 4791 AATGAGATCAAGATCAAGATGTCAAAGCTGAAAGACTAAAGGTCAAAGCTGATGGGT 4850  
Db 3781 AATGAGATCAAGATCAAGATGTCAAAGCTGAAAGACTGAAAGGTCAAAGCTGATGGGT 3840  
Qy 4851 TGGCAGTGTGAGAGAGAAACCATCCATTTATGCTGAGCAATTTGCAATAAGAAATATGAT 4910  
Db 3841 TGGCAGTGTGAGAGAGAAACCATCCATTTATGCTGAGCAATTTGCAATAAGAAATATGAT 3900  
Qy 4911 GACAGAAAGATTTTCTTCTTTCAGAAAGTAAAGAAAGTGGCAACTTAAATACATCTCT 4970  
Db 3901 GACAGAAAGATTTTCTTCTTTCAGAAAGTAAAGAAAGTGGCAACTTAAATACATCTCT 3960  
Qy 4971 TTCTGTGTGAAAAAGGAGAGATCTTTAGGACTATTTGGGCTCAAATGGTGTCTGCAAAAGC 5030  
Db 3961 TTCTGTGTGAAAAAGGAGAGATCTTTGGGACTGTTGGGCTCAAATGGGACTGCAAAAGC 4020  
Qy 5031 ACAATTTATTAATTTCTGTTTGTGTATTTGAAACCAATTTTCAAGCAGCAGGATTTTATGGA 5090  
Db 4021 ACAGTCAATTAACACTCTGTTTGGGCGAGCTTTGAGCCAACTTTCAGGCAAGATCTTTCTTAGGA 4080

Qy	5091	GATTATTTCTTTCAGAGACAACTGAAGATGATGATTTCACTGAAGTGTATGGTTACTGCTCT	5150
Db	4081	GATTATGTTTTCATTTCACTTCAAGTGAAGATGACATCCATTAAGTGTATGGGTACTGCCCC	4140
Qy	5151	CAGATAAACCTTTTGTGGCCAGATATCTATCTGAGGAACTTTTGAATTTATGAGCT	5210
Db	4141	CAGATAAACCTTGTGGCCAGATCTCTCTCTAGGAACTTTTGAATTTATGAGCT	4200
Qy	5211	GTCAAGGAATGAGTCAAGTGCATCAAGAAAGTCTATAGTGCATTAACACATGCACTT	5270
Db	4201	GTGAAGGAATGAGTCTGTGTGCATCAAGAAAGTCTATAGTGCATTAACAAAGCCCTT	4260
Qy	5271	GATTAAAGAACTCTTCAAGAGACTGTAAAGAACTCTCGAGGAATCAACAGAAAG	5330
Db	4261	GATTGAAGAACTCTTCAAGAGACTGTAAAGAACTCTCGAGGAATCAACAGAAAG	4320
Qy	5331	TTGTGTTTCTCTAAGTATGCTAGGGAATCTCTCAGATTTACTTTGTAGTAAGAACCATCT	5390
Db	4321	CTGTGTTTCTCTCAGATGCTGGGGAACCTTCAGTGACGCTGTGGATGAGCGCTCC	4380
Qy	5391	ACAGTATGATCCCAAGCCCAACAGCACATGTGGCGAGCAATTCGAATGCAATTTAA	5450
Db	4381	ACAGGCAATGACCCCAAGCCCAACAGCACATGTGGGAGCTATTCGAATGCAATTTAA	4440
Qy	5451	AAAGAAAGCGGGCTGTATTTGACCACTTCTATATGAGGAGGAGGAGGCTGTCTGT	5510
Db	4441	AAAGAAAGCGGGCGGCTCTCTGACTTCTCTATATGAGGAGGAGGAGGCTGTCTGT	4500
Qy	5511	GATGAGTATGATCATGTGTCTGGGAGCTTATGATGATGATGATGATGATGATGATGAT	5570
Db	4501	GACAGAGTGGCCATCATGTGTCTGGGAGCTTATGATGATGATGATGATGATGATGAT	4560
Qy	5571	AAGATTAATTTGGAAGAGCTTCTTTTGGAAATTAATTTGAAGGAGCTGATGAGAAAC	5630
Db	4561	AAGATTAATTTGGAAGAGCTTCTTTTGGAAATTAATTTGAAGGAGCTGATGAGAAAC	4620
Qy	5631	CTAGAGTATGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCAAGCGCTCAG	5690
Db	4621	CTGGAATATGATCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCAAGCGCTCAG	4680
Qy	5691	GAAGATTTTCTCTATTTTGGCTTATAAATTTCTTAAGGAAGATTTTCACTGCTCTTCA	5750
Db	4681	GAGAGCTTTTCTCTATTTTGGCTTATAAATTTCCAAAGAGATTTTCACTGCTCTTCA	4740
Qy	5751	CAATCTTTTCTTAAAGTGAAGAGCTTAAATGCTTTTCCCAATGCAAGCGCTCAG	5810
Db	4741	CAATCTTTCTGCTAAGTTGGAAGAGCTTAAATGCTTTTCCCAATGCAAGCGCTCAG	4800
Qy	5811	TCTCAAGCAATTTGGAAGAGCTTAAATGCTTAAAGAGAGAGGAGGAGAGAT	5870
Db	4801	TCTCAAGCAATTTGGAAGAGCTTAAATGCTTAAAGAGAGAGGAGGAGAGAT	4860
Qy	5871	AATAGTTTGTGAATTTTAAACAGCACATTTTGTGGGAACGAAACACAAAGAGATAGATA	5930
Db	4861	AACAGCTGTGGACTTTTGGCCAGCACTCTCTGTTGGGAGGAGGAGGAGGAGATAGATA	4920
Qy	5931	GTATTTTGA 5939	
Db	4921	GTATTTTGA 4929	

RESULT 14  
RNO50165  
LOCUS  
DEFINITION Rattus norvegicus mRNA for ATP-binding cassette protein 5 (Abca5 gene), splice variant V+16.  
VERSION AJ50165  
KEYWORDS Abca5 gene; alternative splicing; ATP-binding cassette protein 5.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1	Petry, F. and Hirsch-Ernst, K.I.	Rattus.
AUTHORS	Identification of novel highly conserved splice variants of rat	
TITLE	Abca5	
JOURNAL	Unpublished	
REFERENCE 2	(bases 1 to 4945)	
AUTHORS	Petry, F.	
TITLE	Direct Submission	
JOURNAL	Submitted (17-MAR-2003) Petry F., Toxicology, University of Goettingen, Robert-Koch-Strasse 40, Goettingen, Lower Saxony, D-37075, GERMANY	
FEATURES	Location/Qualifiers	
source	1..4945	
	/organism="Rattus norvegicus"	
	/mol_type="mRNA"	
	/strain="Wistar"	
	/db_xref="taxon:10116"	
	/tissue_type="testis"	
gene	1..4945	
	/gene="Abca5"	
CDS	9..2786	
	/gene="Abca5"	
	/note="splice variant V+16"	
	/codon_start=1	
	/evidence=experimental	
	/product="ATP-binding cassette protein 5"	
	/protein_id="CAD80052.1"	
	/db_xref="GI:29170396"	
	/translation="MATAIRDVGWVRQTRTLLKNVLCRTKSSQVLEPLPLFLP WLIIISMHPNKKYEVDSDIELSPMDKSIILNLIIGTYFTVNTTSSVMQVSTDLHPD VLIVTEYASEKELLASLSKPSFVGVFKDVMYSYELRFFPDMVPSVYMDSRAGCS KSCDAAYWSSGFTALQASIDAAIIQLATNVSRLWRELESTKAVIMGEAAVVIDTFPR GVILYIVAPSPGFLAHIVAEKEKLEKFLIMGLHDTAFWLSVLLVTSILFL MSLMVAIATASSLPQSSSIVIFLLFFYGLSSVFPFALMLTLPFKSKHVGVEFFV TVFVFGLLIVLESFPRSLVWLFSPICQCAFLIGIAQVHLEDFNEGALFSSLTGE PYPLITITMLALDSVFYALLALVDQVPGFELGRSSLYFLKPSYKRNKRYKEL SEGNINGNISLNEIVEPVSFIFGKEAIRISIGIQKAYKQNETVEALRNLSIDYIEGQ ITALLHSGTGTSLMNLICGLCPDPSFASIDGFRVSEIDEMFEARWMIIGICPDSDM NEDVLTVRENSILASVKGIPANNIIQEVOKVLLDLMQAIKDNOKKLSGGQKRLS LGIATVGNFKILLDEPTAGMDPCSRHIVWNLKYRKANRVTVFSTHMDDEALILADR KAVISQGMKLVGSSIFLAKWGIYRUSMIDRYKATESLSLVRQHPHIAALIQON DQIIVYSLFPKDMDFSGLSALDTHSLNIGVISYQVSMTLEDFVLEKLEVEABIDQAD YSVFTQPREBETSKSDMEQESLILSETKASIVSTMSLWQVSTIAKPHFLSLK RESKSVRSVLLLLLIFFAVQIFMFLVHHSFKNAVPIKLVLPDLVFLKPGDKPHKYKTS LLLQNSTGESV"	
BASE COUNT	1395 a 996 c 1101 g 1453 t	
ORIGIN		

Query Match	58.5%;	Score 3816.2;	DB 10;	Length 4945;
Best Local Similarity	85.9%;	Pred. No. 0;		
Matches 4249;	Conservative 0;	Mismatches 679;	Indels 16;	Gaps 1;
Qy	1004	AGAAACATGTCCTCACTGCAATTTAGGAGGTAGGAGTTGGAGACAGACAGCAACTTCT	1063	
Db	2	AGAAACATGTCCTCACTGCAATTTAGGAGGTAGGAGTTGGAGACAGACAGCAACTTCT	61	
Qy	1064	ACTGAAGATTTACTTAATTAATGAGACACAAAGAGTAGTGTTCAGGAATTTCTTTT	1123	
Db	62	TCTGAAGATTTACTTACCTCGTCAGGAGTAAAAGAGTAGTGTTCAGGAATTTCTTTT	121	
Qy	1124	TCCACTATTTTATTTTGGTAAATTAATTAATAGCATGATGATGATGATGATGATGATGAT	1183	
Db	122	TCTCTATTTTCTTATTTTGGCTGATATTAATTAATAGCATGATGATGATGATGATGATGAT	181	
Qy	1184	TGAAGAGTGCCTAATATAGAACTCAATCTATGAGCAAGTTTACTCTTTCTTAATCTAAT	1243	
Db	182	TGAAGAGTATCTGATATAGAACTCAGCCCTATGAGCAAAATCCATCTTTTCAACCTTAT	241	
Qy	1244	TCTTGGATATATCCAGTCACTAATATTAACAGAGCAGCATGATGAGAAAGTGTCTACTGA	1303	
Db	242	CCTGGGGTACACTCCAGTGACTAATACTACAAGCAGCGGTGATGAGAGAGTTTCTTACCGA	301	

Qy	1304	TCATCTACGTGATGTCATTAATTA	CTGAAGAATATACAAATGAAAGAAATGTTAAACATC	1363
Db	302	TCATCTTCCGACCTTCTAGTTACT	GAAGAATACGCCAGTGAGAAGGAGCTGCTGGCATC	361
Qy	1364	CAGTCTCTTAAGCCGAGCACTTTG	TAGTGTGGTTTTCAAGACTTCCATGTCCCTATGA	1423
Db	362	CAGTTTCTTAAGCCGAGCACTTTG	TAGTGTGGTTTTCAAGACTTCCATGTCCCTATGA	421
Qy	1424	ACTTCGTTTTTTTCTCGATATGAT	TCCAGTATCTTCTATTTATATGATTTCAAGAGCTGG	1483
Db	422	ACTCGGTTTTTCTCGATATGAT	TCCAGTATCTTCTATTTATATGATTTCAAGAGCAGG	481
Qy	1484	CTGTTCAAAATCATGTGAGGCTG	CTCAGTACTGTGCTCCTCAGGTTTCAAGTTTTTACAAGC	1543
Db	482	CTGTTCAAAATCATGTGAGGCTG	CTCAGTACTGTGCTCCTCAGGTTTCAAGTTTTTACAAGC	541
Qy	1544	ATCCATAGATGCTGCCATTTAC	AGTTTGAAGCAATGTTTCTTTTGGAGGAGCTGGA	1603
Db	542	CTCAATAGATGCTGCCATTTAC	AGTTTGAAGCAATGTTTCTTTTGGAGGAGCTGGA	601
Qy	1604	GTCAACTTAAAGCTGTTATTTAT	TGGAGAAACTGCTGTTGTAGAAATAGATACCTTTCCCGG	1663
Db	602	GTCAACCAAGCTGTCTATCAT	TGSGGGAAGCTGCGGTCTGTGGAGTTGACACCTTCCACG	661
Qy	1664	AGGAGTAAATTTAATATACCT	TAGTTATAGCATTTTCACTTTTGGATPACTTTTGGCAAT	1723
Db	662	AGGGGTCACTCTCATCTAC	CTCGTATAGCTTCTCGCCCTTCTGGGTACTTTCTGGCAAT	721
Qy	1724	TCATATCGTAGCAGAAAGAA	AAAAAATAAAGAAATTTTAAAGATATGGACCTTCA	1783
Db	722	TCATATCGTAGCAGAAAGAA	AAAAAATAAAGAGGTTTTTAAAGATATATGGACCTTCA	781
Qy	1784	TGATACTGCTTTTGGCTTCT	CTGGTCTCTATATACAAAGTTTAAATTTTCTTATGTC	1843
Db	782	CGACACTGCTTTTGGCTGCT	CTGGGTTCTTCTGTACACAGATTTGATCTTCTTATGTC	841
Qy	1844	CCTTCTTTATGGCAGTCA	TTCGACAGCTTCTTTGTTATTTTCTCAAAAGTAGCAGATGT	1903
Db	842	CCTTCTTTATGGCAGTCA	TTCGACAGCTTCTTTGTTATTTTCTCAAAAGTAGCAGATGT	901
Qy	1904	GATATTTCTGTTTTTCTTT	TATGATATCATCTGTAATTTTTCGCTTTTAAATGCTGAC	1963
Db	902	GATTTTTCTGCTCTTCT	CTGTACCGAATGTCTATCTGTGTTTTTGTCTTTTAAATGTTGAC	961
Qy	1964	ACCTCTTTTAAAAAATCA	AAAAACATGSGGAATAGTTGAAATTTTTCGTACTGTGGCTTT	2023
Db	962	GCCTCTTTTAAAAAATCA	AAAAACATGSGGAGTGGTGTGTTTTTTCACCGTGGTGT	1021
Qy	2024	TGGATTTTATTTGGCCTTA	TATGATATCTCTATAGAAAGTTTTTCCCAATTCGTTAGTGTGGCT	2083
Db	1022	TGGCTTTTGTGGCCTGTT	GATAGTCTCTGTAGAAAGTTTTTCCCAAGTTCGCTAGTGTGGCT	1081
Qy	2084	TTTCAGTCCCTTCTGTCA	CTGTACTTTTGTGATTTGGATTTGGCAGGTCATGCAATTTAGA	2143
Db	1082	CTTCAGTCCCTTGTGT	CAGTGGCTTTCTGATTTGGGATTTGCAAGGTCATGCAATTTAGA	1141
Qy	2144	AGATTTTAAATGAAGTCT	CTCATTTTCAAAATTTGACCTGCAAGGCCCATCTCTCTAATAT	2203
Db	1142	AGATTTCAATGAAGGAG	CCCTTAITTTTCTAGTTTGTACTGAAAGGCCCTTACCCGCTAATAT	1201
Qy	2204	TACAAATATCATGTCA	CACCTTAATAGTATATTTCTATGTCTCTTGTGCTGCTATCTTGA	2263
Db	1202	CACCTTACCATGCTCG	CTCTGGACAGTGTGTTCTATGCTCTGCTGCTGTGTACCTTGA	1261
Qy	2264	TCAGTCAATTCAGGGG	AAATTTTGGCTTACGAGATCATCTTTATATTTTCTGAAGCCTTC	2323
Db	1262	TCAGGTCATCCAGGG	AAATTTTGGCTTACGAGGAGTCACTTTTGTACTTCTCTAAAGCCATC	1321
Qy	2324	ATATTTGTCAAGAGT	AAAGAAATTTAGGAGTATACAGGGCAATCTTTAATGGAAA	2383
Db	1322	ATATTTGTCAAAAAC	AAAAAGGAACATAAGGAGCTATACAGGGGCAACATCAATGGGAA	1381
Qy	2384	TATTAGTTTTTAGTGA	AAATTTATGAGCCAGTTTCTTCTCAGAAATTTGTAGAAAAGAGCCAT	2443

Db	1382	TATTAGTCTCAATGAAATTTG	TGAGCCCGTTTTCTTCAGAAATTTATAGAAAAGAGCTAT	1441
Qy	2444	AGCAATTTAGTGTGATTTT	CAGAAGACATACAGAAAGAGGGTGAATAATCTCGAGGCTTTGAG	2503
Db	1442	AGAAATAAGTGTGATTTT	CAGAAAGCCCTATAGAAAGAAATGAGACCGTGGAGGCTTTGAG	1501
Qy	2504	AAATTTGTCAATTTGAC	ATATATGAGGGTTCAGATTTACTTGCCTTACTTTGGCCACACAGTGAAC	2563
Db	1502	AAATTTGTGCTTTGAC	ATCTATGAAGCCAGATTTACTTGTCTCTGCTGCGCCACACAGTGAAC	1561
Qy	2564	AGAAAGAGTACATTTGAT	GAATATTTCTTTGTGAGCTCTGCCACCTTCTGATGGGTTTGC	2623
Db	1562	AGAAAAAAGCACACTA	ATGAATTTCTGTGTGAGCTGTGCCACCTTCTGATGGGTTTGC	1621
Qy	2624	ATCTATATATGACAC	CAGAGTCTCAGAAATAGATGAATAATGTTTGAACAGAGAAATATGAT	2683
Db	1622	TTCTATATATGACAC	CAGAGTCTCTGAAATAGATGAATAATGTTTGAACAGAGAAATATGAT	1681
Qy	2684	TGGCATTTTGTCCA	CAGTTAGATATACATTTTGTATGTTTGTGACAGTGAAGAAAATTTATC	2743
Db	1682	TGGCATATTTTCCA	CAGTTAGATATGAACTTTTGTATGTTTGTGACAGTGAAGAAAATTTATC	1741
Qy	2744	AATTTTGGCTTCAAT	CAAGGGATACAGCCAAATATAATAACAAGAGTGCAGAAAGT	2803
Db	1742	AATTTTGGCTTCAAT	CAAGGGATACAGCCAAATATAATAACAAGAGTGCAGAAAGT	1801
Qy	2804	TTTACTAGATTTAG	CATGACAGTATCAAGATTAACCAAGCTTAAAAAATTAAGTGTG	2863
Db	1802	TTTACTAGATTTAG	CATGACAGTATCAAGATTAACCAAGCTTAAAAAATTAAGTGTG	1861
Qy	2864	TCAAAAAAGAAAGCT	GTCTATTAGGAATTTGCTTTTGGGAACCCCAAGATATCTGCTGCT	2923
Db	1862	CAAAAAGAGAAAGCT	GTCTATTAGGAATTTGCAAGTCTTTGGGAATTCGAAGATCTCTGCTGCT	1921
Qy	2924	AGATGAACCAACAG	CTGGAATGGAACCCCTGTTCTCGACATATTTGTATGGAATCTTTTAAA	2983
Db	1922	AGATGAGCGACAG	CAGGAATGGAACCCCTGCTCTCGCCACATTTGTTGGAATCTTCTCAA	1981
Qy	2984	ATACAGAAAGCCAA	TGCGGTGACAGTGTTCAGTACTCATTTTCATGATGAAGTGCAT	3043
Db	1982	GTACAGAAAGGCT	TAAACCGAGTGCAGTCTTCAGCACTCACTTCATGATGAGGCCGACAT	2041
Qy	3044	TCCTTGAGATAG	AAAAGCTGTGTATATCAAGAGAAATGCTGAAATGTTGCTGCTTCAAT	3103
Db	2042	TCCTTGCGGACAG	AAAGCTGTGTATATCAAGAGAAATGCTGGAATGTTGTTGTTCTTCAAT	2101
Qy	3104	GTTTCTCTCAAAAG	TAAATGSGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATG	3163
Db	2102	TTTTCTGAAAAG	TAAATGSGGAATCGGCTACCGCTGAGCATGTATATAGACAGGTA	2161
Qy	3164	TGCGACAGAAAT	CTCTTTCTTCTTCTACTGTTTAAACAAATATACCTGGAGCTACTTTTATACA	3223
Db	2162	TGCGACAGAGT	CGCTCTCGTGTGTTAGACAGCACTCCCTGCGGCGCTCTGCTGCA	2221
Qy	3224	ACAGAAATGACCA	CAACTTGTGTATAGTCTTGCCCTTCAAGGACATGACAAAATTTTTCAGG	3283
Db	2222	GCAGAAATGACCA	CAAGTGTCTATAGCTTGGCCCTTCAAGAGCATGGAACAAATTTTTCAGG	2281
Qy	3284	TTTGTGTTTCTG	CCCTAGACAGTCAATTTCAAAATTTTGGGTGTCAATTTCTTATGGGTGTTCCAT	3343
Db	2282	CTTGTGTTTCTG	CTCTAGACATCAATTTCAAACTGGGTGTAAATTTCTTATGTTGTTCCAT	2341
Qy	3344	GACGACTTTGGA	AGAGCTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGA	3403
Db	2342	GACGACTTTGGA	AGAGCTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGA	2401
Qy	3404	TTATAGTGTATTT	TACTCAGCAGCTCTGGAGGAAGAAATGGAATTTCAAAATCTTTTGTATGA	3463
Db	2402	TTATAGTGTATTT	TACCAGAGCCCGGAGGGAAGAAACAGATTTCAAGATCTTTTGTATGA	2461
Qy	3464	AATGGAACAGAG	CTTACTTATTTCTTCTGAAAACCAAGGCTTCTCTAGTGTGAGCCCATGAG	3523

2462 AATGGAACAGAGTTTACTTATCTCTCTGAAACCAAAAGCTTCTCTAGTGAGCAACCATGAG 2521 Db  
3524 CCTTTGGAAACAACAGATGATATCAATAGCAAAAGTTTCAATTTCTTTACCTTGAACCGTGA 3583 Qy  
2522 TCTCTGGAAGCAGCAAGTATCTACAATAGCAAAAGTTTCAATTTCTCTCACTGAAGCGAGA 2581 Db  
3584 AAGTAATCAGTCAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTT 3643 Qy  
2582 AAGCAATCAGTCGCTCTGCTGCTTCTGCTTTTAAATTTTTCAGTTCAGATTTT 2641 Db  
3644 TATGTTTTGGTTCATCACTCTCTTTTAAATAGCTGTGGTTCCCAATCAAACTTGTTCAGA 3703 Qy  
2642 TATGTTTTGGTTCATCACTCTTTTAAATAGCTGTGGTTCCCAATCAAACTTGTTCAGA 2701 Db  
3704 CTTATATTTTCTAAACCTCGAGACAAACACATAAATAAACAACAGTCTGCTCTTCA 3763 Qy  
2702 CTTGTATTTTCTAAACCTCGAGATAAACCTCATAAATAAACAACAGCCTGCTGCTTCA 2761 Db  
3764 AAATTTCTGCT-----GACTCAGATATCAGTGATCTTATTAGCTTTTCA 3807 Qy  
2762 AAATTTCTAGGTCAGAGCGTGTGAAGATTCAGATATCAATGATCTTATTGACTTTTCA 2821 Db  
3808 CAAGCCAGAACATAATCGTGACGATGATTAATGACAGTGAATGATATCGTGCTGCTCCC 3867 Qy  
2822 CACAGCAGAACATAATTTGTGCCATGTTTAAATGACAGTGAATGATGCTGCTGCTCCTC 2881 Db  
3868 ATAGTCGGCTTTAAATGATGATTCATTCAGAAAGGACTATGTTTGTGACGCTGTTTCA 3927 Qy  
2882 ATAGTCGGCTCTCAATGTGGTGAGTCAGAAAGGACTATGTTTGTGCTGCTGTTTCA 2941 Db  
3928 ACAGTACTATGCTTTTACTTATATAGTGAATATCATTTAGTAATCTACTATCTTT 3987 Qy  
2942 ACAGTACTATGCTTTTACTTATGCTGAGTATGATGATGATGATGATGATGATGATGAT 3001 Db  
3988 ATCAATTAATGAGTGAACCAATCCAGATCTGGAGTACCCATCTTTTCAAGAAATTA 4047 Qy  
3002 ATCAATTAATGAGTGAACCAATCCAGATCTGGAGTACCCGCTTCATTCAGAGAAATTA 3061 Db  
4048 CTGATATAGTTTTTAAATGAGTGAATTTTCAAGCAGCTTGTGCTGGAATCATTTGTTA 4107 Qy  
3062 CTGACATGTTTTTAAAGTTGAGCTATATTTTCAAGCAGCTTGTGCTGGAATCATTTGTTA 3121 Db  
4108 CTGCAATGCCACCTTACTTTGCCATGGAATGAGAGATCATATAAGATCAAAAGCTTATA 4167 Qy  
3122 CTGCAATGCCACCTTACTTTGCCATGGAATGAGAGATCATATAAGATCAAAAGCTTATA 3181 Db  
4168 CTCAACTTAACTTTTCAGGCTTTTGGCATCTGCGATATGGAATGGACAAGCTGTGTTG 4227 Qy  
3182 CTCAGCTTAAACTTTTCAGGCTTTTGGCATCTGCGATCTGGAATGGACAAGCTGTGTTG 3241 Db  
4228 ATATCCCTTATTTTATCATTTCTTATTTGATGCTAGGAAGCTTACTGGCAATTCATTT 4287 Qy  
3242 ACATTTCCCTGTTTTTGTGTTCTGATCTTGTGCTGGAAGTTTATTTGCTTTTCATC 3301 Db  
4288 ATGATTTATATTTTATCTGTAAGTTCTGCTGCTGTTTTTGTGCTTTTATTTGTTATG 4347 Qy  
3302 ACGACTGTATTTTATCTGTAAGTTCTGCTGCTGTTTTTGTGCTTTTATTTGCTTTATG 3361 Db  
4348 TTCCATCAGTTATTTCTGTTCACTTATATGCTTTCTTTTCACTTTTAAAGAAATTTTAAATA 4407 Qy  
3362 TGCCCTCGGTATTTCTGTTTATATAGCTTCTGTTTCACTTTTAAAGAAATTTTAAATA 3421 Db  
4408 CCAAGAAATTTTGTGCTATTTCTATTTGATGCTAGGAAGCTTACTGGCAATTCATTT 4467 Qy  
3422 CCAAGAAATTTTGTGCTATTTCTATTTCTGTCAGCAGTTGCTGTTGCTGCTGCTGCTG 3481 Db  
4468 AAATAAATTTTCTGTTGGAATACAAATGCAACTATTTCTTCAATTTGCTTTTGTATCA 4527 Qy  
3482 AAATAAATTTTCTGTTGGAATACAAATGCAACTATTTCTTCAATTTGCTTTTGTATCA 3541 Db  
4528 TCATTTCAATTTCTATCTTCTAGGTTGCTGATTTCTTTTCAATAAGATTTCTTTGGAAGA 4587 Qy  
3542 CCATCCCAATCTACCCGCTCTCTGCTGTTGCTGATTTCTTTTCAATAAGGGTTCTTTGGAAGA 3601 Db

4588 ATGTACGAAAAAATGTGGACACCTATAATCCATGGGATAGGCTTTTCACTAGCTGTTATAT 4647 Qy  
3602 ATATACCAAAATCAGAAATGCTTACAAATCCCTGGGACAGACTTTTGTAGTTGCTGTAATCA 3661 Db  
4648 CGCCTTACCTGAGTGTGTACTGTGGATTTTCTCTTACAATACTATGAGAAAAAATATG 4707 Qy  
3662 TGCCCTTACCTGAGTGTGTACTGTGGATTTTCTCTTACAACACTATGAGAAAAAATATG 3721 Db  
4708 GAGGAGATCAATAGAAAAAGATCCCTTTTTCAGAAAACTTTTCAACGAAGCTTAAAAATA 4767 Qy  
3722 GAGGAGATCAATAGAAAAAGATCCCTTTTTCAGGGCCCTCTCACAATAAGCCAAACATA 3781 Db  
4768 GGAAGCTTCCAGAACCCAGACAATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGAC 4827 Qy  
3782 AGAAGTTTCCAGAACCCAGACAATGAGGATGAAGATGAAGATGTCAAAGCGGAAAGGT 3841 Db  
4828 TAAAGTCAAGAGCTGATGGGTTGCCAGTGTGTTGTGAGAGAAACCATCTATTATGTTCA 4887 Qy  
3842 TGAAGTTTAAAGAGCTGATGGGTTGCCAGTGTGTTGTGAGGAGAAACCGGCCATTATGTTAT 3901 Db  
4888 GCAATTTGCATAAAGAAATATGATGACAAGAAAGATTTTCTTTTCAAGAAAGTAAAGA 4947 Qy  
3902 ACAATTTGCATAAAGAAATATGATGACAAGAAAGATTTTCTTTTCACTCAAGGAAACAAACA 3961 Db  
4948 AAGTGGCAACTAAATAACATCTCTTTCTGTGTGAAAAAGGAGAGATCTTTAGGACTATTGG 5007 Qy  
3962 AAGTAGCAACAAATAACGTCCTTTCTGTGTGAAAAAGGAGAGATCTTTGGGACTGTTGG 4021 Db  
5008 GTCCAAATGTGCTGGCAAAAGCAAAATTAATTAATTTCTGTTGTTGTTGATTAATGAACAA 5067 Qy  
4022 GTCCAAATGTGCTGGCAAAAGCAGATTAATTAATTTCTTTGTTGGGATGTTGAACCAA 4081 Db  
5068 CTTTCAAGCCAGATTTTGTAGGAGATTTCTTTCAGAGACAGTGAAGATGATGATTCAC 5127 Qy  
4082 CTTTCAAGCCAGATTTTGTAGGAGATTTCTTTCAGAGACAGTGAAGATGATGATTCAC 4141 Db  
5128 TGAAGTGTATGGTTTACTGCTCTCAGATAAAACCCCTTTGTGGCCAGATATCACTATGTCAG 5187 Qy  
4142 CCAAGTGTATGGTTTACTGCTCTCAGATAAAACCCCTTTGTGGCCAGATATCACTATGTCAG 4201 Db  
5188 AACATTTTGAATTTATGAGCTGTCAAAGGAATGAGTGCAGATGACATGAAGAAATCA 5247 Qy  
4202 AACATTTTGAATTTATGAGCTGTCAAAGGAATGAGTGCAGATGACATGAAGAAATCA 4261 Db  
5248 TAAAGTCGATACATGACATTTGATTTAAAGAACATCTTCAGAGACAGTGAAGAAATCA 5307 Qy  
4262 TCAGTCGAATAAAGAGCCCTTGAATTTGAAAGAACATCTTCAGAGAACCTGTAAGAAATCA 4321 Db  
5308 TACCTGCAAGAAATCAAAAGAAAGTTGTGTTTGTCTTAAGTATGCTAGGAAATCTCTCAGA 5367 Qy  
4322 TCCCTGCAAGAAATCAAAAGAAAGTTGTGTTTGTCTTAAGTATGCTAGGAAATCTCTCAGA 4381 Db  
5368 TTAATTTGCTAGATGAACCATCTCAGGTATGGAATCCCAAGCCAAACAGACATGTTGGC 5427 Qy  
4382 TGACACTGCTGATGAGCCGTCTCAGGTATGGAATCCCAAGCCAAACAGACATGTTGGC 4441 Db  
5428 GAGCAATTCGAATTCATTTTAAAGAACAGAAAGGCTGCTTATCTTGACCACTCACTATA 5487 Qy  
4442 GAGCAATTCGAATTCATTTTAAAGAACAGAAAGGCTGCTTATCTTGACCACTCACTATA 4501 Db  
5488 TGGAGGAGCAGAGCTGCTGTTGATCGAGTATGATCATGTTGCTGCTGCTGCTGCTGCTGCT 5547 Qy  
4502 TGAAGAGCGGAGGCTGCTGTTGATCAGGGTGGCCATCATGTTGCTGCTGCTGCTGCTGCTG 4561 Db  
5548 GTATCGGAACAGTACAACTCTAAAGAGTAAATTTGGAAGAAAGCTACTTTTTTGAATAATTA 5607 Qy  
4562 GCATTTGGAACAGTACAACTCTAAAGAGTAAATTTGGAAGAAAGCTACTTTTTTGAATAATTA 4621 Db  
5608 AATTGAGGAGCTGATGAGAAACCTTAGAGTAGAGCGCCCTTCAAGAGAAATTCAGTATA 5667 Qy  
4622 AACTAAGAGCTGATGAGAAACCTTGAAGTATGATGCTGCTTCAAGAGAAATTCAGTATA 4681 Db





4063 AAATTGAGCTGTATTTTCAAGCAGCTTTGCTTGAATCAATTTGTTACTGCAATGCCACCTT 4122  
1081 AAATTGAGCTGTATTTTCAAGCAGCTTTGCTTGAATCAATTTGTTACTGCAATGCCACCTT 1140  
4123 ACTTTGCGCATGGAAAAATGCGAGAAATCATAGATCAAAAGCTTATATCTCAACTTTAAACTTT 4182  
1141 ACTTTGCGCATGGAAAAATGCGAGAAATCAATA----- 1170  
4183 CAGGTCTTTTGCCATCTGCATATTTGATTTGGACAAAGCTGTGTTGATATCCCTCTATTTT 4242  
1171 -AGGTCTTTTGCCATCTGCATATTTGATTTGGACAAAGCTGTGTTGATATCCCTCTATTTT 1229  
4243 TTATCATTTCTATTTTATCTAGCTAGGAAGCTTACTGCGCAATTTCAATATGATGATATATTTT 4302  
1230 TTATCATTTCTATTTTATCTAGCTAGGAAGCTTACTGCGCAATTTCAATATGATGATATATTTT 1289  
4303 ATACTGTAAAGTTCCCTTGCTGTGGTTTTTTCCTTATTTGGTTATGTTCCATCAAGTTATTC 4362  
1290 ATACTGTAAAGTTCCCTTGCTGTGGTTTTTTCCTTATTTGGTTATGTTCCATCAAGTTATTC 1349  
4363 TGTTCACTTATATGCTTCTTTTCACTTTAAGAAAATTTTAAATACCAAGAAATTTTGGT 4422  
1350 TGTTCACTTATATGCTTCTTTTCACTTTAAGAAAATTTTAAATACCAAGAAATTTTGGT 1409  
4423 CATTTATCTATTTCTGTGGCAGCTTGNCTTGTATTTGCAATCACTGAAATAAATCTTCTTTA 4482  
1410 CATTTATCTATTTCTGTGGCAGCTTGGCTTGTATTTGCAATCACTGAATTAACCTTCTTTA 1469  
4483 TGGGATACACAAATTTGCAATTAATTTCTTCAATATGCTTTTGTATCATCATTTCCAATCTATC 4542  
1470 TGGGATACACAAATTTGCAATTAATTTCTTCAATATGCTTTTGTATCATCATTTCCAATCTATC 1529  
4543 CACTTCTAGCTTGCCTGATTTCTTTTCAATAAAGATTTCTTGGAGAATGTACGAAAAATG 4602  
1530 CACTTCTAGCTTGCCTGATTTCTTTTCAATAAAGATTTCTTGGAGAATGTACGAAAAATG 1589  
4603 TGGACACCTAATPCCATGGATAGGCTTTTCACTAGCTGTATATCGCTTTACTCTGAGT 4662  
1590 TGGACACCTAATPCCATGGATAGGCTTTTCACTAGCTGTATATCGCTTTACTCTGAGT 1649  
4663 GTGTACTGTGGATTTTCTTCAATPACTATGAGAAAAATATGGAGGCGAGATCAATAA 4722  
1650 GTGTACTGTGGATTTTCTTCAATPACTATGAGAAAAATATGGAGGCGAGATCAATAA 1709  
4723 GAAAAAGTCCCTTTTTCAGAAAACCTTTCAACGAAGTCTAAATATGGAAGCTTCCGAAC 4782  
1710 GAAAAAGTCCCTTTTTCAGAAAACCTTTTCAACGAAGTCTAAATATGGAAGCTTCCGAAC 1769  
4783 CACCAGACAAATGAGATGAAGATGAAGATGTCAAAGCTGAAAGCTAAAGGTCAAAGAGC 4842  
1770 CACCAGACAAATGAGATGAAGATGAAGATGTCAAAGCTGAAAGCTAAAGGTCAAAGAGC 1829  
4843 TGATGGGTGCGAGTGTGTGAGGAGAAACCATTCATTTGTTGTCAGCAATTTGCGATAAAG 4902  
1830 TGATGGGTGCGAGTGTGTGAGGAGAAACCATTCATTTGTTGTCAGCAATTTGCGATAAAG 1889  
4903 AATATGATGACAGAAAGATTTTCTTTCAGAAAAGTAAAGAAAGTGCGCAACTAAAT 4962  
1890 AATATGATGACAGAAAGATTTTCTTTCAGAAAAGTAAAGAAAGTGCGCAACTAAAT 1949  
4963 ACATCTCTTTCTGTGTCAGAAAAGGAGAGATCTTGGACATTTTGGGTCCAAATGGTGCTG 5022  
1950 ACATCTCTTTCTGTGTCAGAAAAGGAGAGATCTTGGACATTTTGGGTCCAAATGGTGCTG 2009  
5023 GCAAAAGCACAATTTATTAATTTCTGTTGGTGTGATTTGAAACCAACTTCAGGGCCAGGTAT 5082  
2010 GCAAAAGCACAATTTATTAATTTCTGTTGGTGTGATTTGAAACCAACTTCAGGGCCAGGTAT 2069  
5083 TTTTGGAGATTTATTTCTTCAGAGCAAGAGTGAATGATTTCACTGAAGTGTATGGGTT 5142  
2070 TTTTGGAGATTTATTTCTTCAGAGCAAGAGTGAATGATTTCACTGAAGTGTATGGGTT 2129  
5143 ACTGTCCTCAGATAAACCCCTTTTGTGGCCAGATACTACATTTGCAGGAACATTTTGAATTT 5202

2130 ACTGTCCTCAGATAAACCCCTTTTGTGGCAGATCTACTACATTTGCAGGAACATTTTGAATTT 2189  
5203 ATGGAGCTGTCAAAGGAATGAGTGAAGTGAATGAAGAAAGTCAATAGTCCGAATAACAC 5262  
2190 ATGGAGCTGTCAAAGGAATGAGTGAAGTGAATGAAGAAAGTCAATAGTCCGAATAACAC 2249  
5263 ATGCACCTTGATTTTAAAGAACATCTTTCAGAGACATGTTAAAGAAACTACTCTCGAGGAATCA 5322  
2250 ATGCACCTTGATTTTAAAGAACATCTTTCAGAGACATGTTAAAGAAACTACTCTCGAGGAATCA 2309  
5323 AACGAAAGTGTGTTTCTCTAAGTATGCTAGGAATCTCAGATTTACTTTGCTAGATG 5382  
2310 AACGAAAGTGTGTTTCTCTAAGTATGCTAGGAATCTCAGATTTACTTTGCTAGATG 2369  
5383 AACCATCTACAGGTATGATCCCAAGCCAAACAGCACATGTTGGGAGCAATTCGAACTG 5442  
2370 AACCATCTACAGGTATGATCCCAAGCCAAACAGCACATGTTGGGAGCAATTCGAACTG 2429  
5443 CATTTAAAAACAGAAAGCGGCTGCTATTTCTGACCACTCACTATATGGAGGAGGAGG 5502  
2430 CATTTAAAAACAGAAAGCGGCTGCTATTTCTGACCACTCACTATATGGAGGAGGAGG 2489  
5503 CTGCTGTGATCGAGTAGCTATCATGTGTCTGGGCACTTAAGATGTATCGGAACAGTAC 5562  
2490 CTGCTGTGATCGAGTAGCTATCATGTGTCTGGGCACTTAAGATGTATCGGAACAGTAC 2549  
5563 AACATCTAAGAGTAAATTTTGGAAAGGCTACTTTTGGAAATTAATTTGAAGGACTCGA 5622  
2550 AACATCTAAGAGTAAATTTTGGAAAGGCTACTTTTGGAAATTAATTTGAAGGACTCGA 2609  
5623 TAGAAAACTAGAACTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAATGCAA 5682  
2610 TAGAAAACTAGAACTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAATGCAA 2669  
5683 GCCGTGAGAAAGTTTCTTCTATTTTGGCTTATAAATTTCTTAAGGAAGATGTTCACT 5742  
2670 GCCGTGAGAAAGTTTCTTCTATTTTGGCTTATAAATTTCTTAAGGAAGATGTTCACT 2729  
5743 CCCTTTCACATCTTTTAAAGCTGGAAGAGCTTAAACATGCTTTTCCCATTTGAAGAT 5802  
2730 CCCTTTCACATCTTTTAAAGCTGGAAGAGCTTAAACATGCTTTTCCCATTTGAAGAT 2789  
5803 ATAGCTTTTCTCAACCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGG 5862  
2790 ATAGCTTTTCTCAACCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGG 2849  
5863 AGGAAGATAATAGTTGTGGAACCTTTTAAACAGCACACTTTTGGTGGGAAACGAAACAGAG 5922  
2850 AGGAAGATAATAGTTGTGGAACCTTTTAAACAGCACACTTTTGGTGGGAAACGAAACAGAG 2909  
5923 ATAGAGTAGTATTTGAATTTGTTTGGTCTGCTTACTGGGACCTTTCTTTCTTTTC 5982  
2910 ATAGAGTAGTATTTGAATTTGTTTGGTCTGCTTACTGGGACCTTTCTTTCTTTTC 2969  
5983 ACTTAATTTAACTTTTGGTTTTTAAAGTTTTTATTTGGAATGCTAACTGGAGAACCAAGA 6042  
2970 ACTTAATTTAACTTTTGGTTTTTAAAGTTTTTATTTGGAATGCTAACTGGAGAACCAAGA 3029  
6043 ACGCACTTGAAATTTTCTAAGCTCCTTAATTTGAAATGCTGTGGTGTGTTGCTTTT 6102  
3030 ACGCACTTGAAATTTTCTAAGCTCCTTAATTTGAAATGCTGTGGTGTGTTGCTTTT 3089  
6103 TCTTTAAATAAACGATATGATTAATTAAGTGAAGCTGCAATGTTTATTTGAAGTATATTG 6162  
3090 TCTTTAAATAAACGATATGATTAATTAAGTGAAGCTGCAATGTTTATTTGAAGTATATTG 3149  
6163 AACTATATAGTTTGTATGTCATCTTTTCCCACTTTCAGAAAACAGTCTTCTGAAATTTG 6222  
3150 AACTATATAGTTTGTATGTCATCTTTTCCCACTTTCAGAAAACAGTCTTCTGAAATTTG 3209  
6223 ATTTAAGGAATTTGTAATAGAAATGTTTTTATTTTAAAGTATCTTTTAAAGTATGCGCATC 6282



Db	3210	ATTAAAGGAATTGTAATAGAAATAGTTTATTTTAAAGTTATCTTTAAGTTTAAGCCATC	3269
Qy	6283	TTCTTAAATAAGTACGTAATGTTCCCAATCTAAATATAAAAACTAATACATAACTAATGCGAT	6342
Db	3270	TTCTTAAATAAGTACGTAATGTTCCCAATCTAATATAAAAACTAATACATAACTAATGCGAT	3329
Qy	6343	AGAAAAGATACATAAAGC	6360
Db	3330	AGAAAAGATACATAAAGC	3347

Search completed: January 31, 2004, 00:03:03  
Job time : 15262 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 19:12:41 ; Search time 282 Seconds  
(without alignments)

10212.854 Million cell updates/sec

Title: US-10-005-338B-1

Perfect score: 6525

Sequence: 1 aaatgtgatattttctt.....ttgatcataataagtgaat 6525

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:\*
- 1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*
  - 2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*
  - 3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*
  - 4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*
  - 5: /cgn2\_6/ptodata/2/ina/PCTUS COMB.seq.\*
  - 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	159.2	2.4	5894	3	US-08-665-259-24
2	159.2	2.4	5894	3	US-08-762-500-24
3	159.2	2.4	6525	3	US-08-762-500-74
4	103.4	1.6	1064	4	US-08-858-207A-88
5	92.2	1.4	5625	4	US-08-961-527-1
6	80.6	1.2	795	4	US-08-961-527-264
7	77.2	1.2	696	4	US-09-134-001C-987
8	75.8	1.2	1026	4	US-09-134-001C-1205
9	75	1.1	774	4	US-09-107-532A-412
10	74.6	1.1	1664976	4	US-08-916-421B-1
11	72.2	1.1	768	4	US-09-107-532A-1681
12	71	1.1	1664976	4	US-08-916-421B-1
13	69.8	1.1	990	4	US-09-107-532A-1190
14	69.2	1.1	969	4	US-09-107-532A-1008
15	68.4	1.0	19124	2	US-08-487-826B-13
16	65.8	1.0	4848	4	US-08-961-527-185
17	65.8	1.0	5717	4	US-08-961-527-102
18	65	1.0	1770	4	US-09-328-352-3466
19	64.6	1.0	3144	4	US-08-961-527-224
20	64.6	1.0	6273	4	US-08-961-527-21
21	63.2	1.0	7218	1	US-08-232-463-14
22	63	1.0	14672	4	US-08-961-527-111
23	62.4	1.0	580073	4	US-08-545-528D-1
24	62	1.0	1230025	4	US-09-198-452A-1
25	61.4	0.9	1797	4	US-09-107-532A-2383
26	59.6	0.9	858	4	US-09-107-532A-2722
27	59.2	0.9	678	4	US-09-134-001C-378

28	58.8	0.9	15567	4	US-09-627-376-3	Sequence 3, Appli
29	58.4	0.9	5852	1	US-07-867-106-2	Sequence 2, Appli
30	57.6	0.9	1038	4	US-09-107-532A-195	Sequence 195, App
31	57.6	0.9	6846	4	US-08-961-527-198	Sequence 198, App
32	57.2	0.9	690	4	US-09-134-001C-1276	Sequence 1276, Ap
33	57	0.9	876	4	US-09-107-532A-551	Sequence 551, App
34	56.2	0.9	873	4	US-09-134-001C-436	Sequence 436, App
35	56.2	0.9	1404	4	US-09-134-001C-615	Sequence 615, App
36	56	0.9	618	4	US-09-107-532A-1929	Sequence 1929, Ap
37	56	0.9	771	4	US-09-107-532A-554	Sequence 554, App
38	55.8	0.9	813	4	US-09-134-001C-2724	Sequence 2724, Ap
39	55.2	0.8	783	4	US-09-107-532A-1996	Sequence 1996, Ap
40	54.8	0.8	933	4	US-09-134-001C-2350	Sequence 2350, Ap
41	54	0.8	765	4	US-09-107-532A-1329	Sequence 1329, Ap
42	54	0.8	7218	1	US-08-232-463-14	Sequence 14, Appl
43	52.8	0.8	747	4	US-09-134-001C-894	Sequence 894, App
44	52.2	0.8	948	4	US-09-328-352-262	Sequence 262, App
45	51.8	0.8	1014	4	US-09-134-001C-2713	Sequence 2713, Ap

ALIGNMENTS

RESULT 1  
US-08-665-259-24  
; Sequence 24, Application US/08665259  
; Patent No. 6028173  
; GENERAL INFORMATION:  
; APPLICANT: Landes, Gregory M.  
; APPLICANT: Burn, Timothy C.  
; APPLICANT: Connors, Timothy D.  
; APPLICANT: Dackowski, William R.  
; APPLICANT: Van Raay, Terence J.  
; APPLICANT: Klinger, Katherine W.  
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES, AND USING SAME  
; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENZYME CORPORATION  
; STREET: One Mountain Road  
; CITY: Framingham  
; STATE: Massachusetts  
; COUNTRY: United States of America  
; ZIP: 01701

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/665,259  
FILING DATE: 17-JUN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Dugan, Deborah A.  
REGISTRATION NUMBER: 37,315  
REFERENCE/DOCKET NUMBER: IGS-9.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508) 872-8400  
TELEFAX: (508) 872-5415  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5894 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2..5053  
US-08-665-259-24

Query Match 2.4%; Score 159.2; DB 3; Length 5894;



2274 CCAGGGGAATTTGGCTTACGAGATCATCTTTATATTTTCTGAAGCCTTCATATTTGTC 2333  
1364 CCAGGGCAGTTCGGGGTCTCAGCCCTGGTACTCTTCTCATGTCCTCTATTTGGTGT 1423  
2334 AAGAGTAAAGAAATATAGAGGATTTATCAGAGGCAATGTTAATGGAATATATAGTTTT 2393  
1424 GGAAGCAGGCGGTTCAGGGAAGGAGGAGAGACAGTACCCCGAGAAAGCACTC 1483  
2394 AGTGAATATTAGCCAGTCTTCTCAGAAATTTGTAGAAAAGAGCATAAGAAATAGT 2453  
1484 AGAAACGAGTACTTTGAGCGGAGCCAGAGGACCTGG---TGGCGGGATCAGATCAAG 1540  
2454 GGTATTACAGAGACATACAGAAAGAGGGTGAATAATGTGAGGCTTTGAGAAATTTGTCA 2513  
1541 CACCTGTCCAAAGTGTTCAGGGTGGGAAATAAGACAGAGGCGCGCTCAGAGACCTGAAC 1600  
2514 TTTGACATATATAGGGGTAGATTACTGCTTACTTGGCCACAGTGGAAACAGAAAGAGT 2573  
1601 CTCAACCTGTACGAGGACAGATCACCGTCTGCTGGGCCACACCGTGCCTGGGAGACC 1660  
2574 ACATTGATGATATCTTTTGTGACTCTGCCACCTCTCTGATGGTGTTCATCTATATAT 2633  
1661 ACCACCTCTCCATGCTCAGAGTCTCTTCCCCCACCAGTGGACGGGCATACATCAGC 1720  
2634 GGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAATGATTGGCATTTGT 2693  
1721 GGGTATGAATTTCCAG-----GACATGTTTCTGATCGGAGAGCCTGGGCGCTGTC 1774  
2694 CCACAGTATGATATACACTTTGATGTTTTCAGAGTGAAGAAAATTTATCAATTTGGCT 2753  
1775 CCGCAGCAGCAGATCTCTGTTTGAACACTTGACAGTTCGACAGCAGCCTTTATTTCTACGCC 1834  
2754 TCAATCAAGGGATACAGCCACCAATATATACAGAGTGCAGAGGTTTACTAGAT 2813  
1835 CAGCTGAAGGCGCTGTCACGTCAAGAGTGCCTTGAAGAGTCAAGCAGATGCTGCACATC 1894  
2814 TTAGACATGACAGACTATCAAGAGTAAACCAAGTCAAGTAAAGTAAAGTAAAGTAAAG 2873  
1895 ATCGGCTGAGGACAGTGAAGTCAAGGAGCGCTTCTGAGCGGGGCGATGAGCGC 1954  
2874 AAGCTGTATAGAAATGCTGTTTGGGAAACCAAGATATCTGCTGCTAGATGAACCA 2933  
1955 AAGCTCTCCATCGCATCGCCTCATCGCAGGCTCAAGGTGCTGATCTGAGCAGGCC 2014  
2934 ACAGCTGAATGACCCCTGTTCTCGACATATGATGGAATCTTTAAATACAGAAA 2993  
2015 ACCTCGGGCATGAGCGCATCTCCAGGAGGCGCATCTGGGATCTTTCTCAGCGGCGAAA 2074  
2994 GCCAATCGGGTGAAGAGTTCAGTACTCATTTATGATGAAGTGAAGTTCATGAGAT 3053  
2075 AGTACCGCCACCATCTGCTGACACCCACTTTCATGACGAGGCTGACCTGCTGGGAGC 2134  
3054 AGAAAGCTGTATATACAGAGGAATGCTGAATGTTGTTTCTCAATGTTCTCAAA 3113  
2135 CGCATGCCATCATGCGCAAGGGGAGCTGCAGTGTGCTGGGCTCTCGCTTCTCTCAAG 2194  
3114 AGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGTCACAGAA 3173  
2195 CAGAAATACGGTGGCGGCTATCATATGCGTGGTGAAGGAGCGGCACTGCAACCCGGAA 2254  
3174 TCTCTTTCTTCACTGGTTAAACACATATACCTGGAGCTACTTTTATTAACAGATGAC 3233  
2255 GACATCTCCAGCTGGTCTCACACCGCTGCCACCGCTGAGAGAGAGCGGCTGGG 2314  
3234 CAACAACTGTGTATAGCTTGCCTTTCAAGGACATGACAAAATTTTTCAGGTTTGTCT 3293  
2315 GCGAGCTGTCTTTCATCTTCCAGAGAGAGCAGCAGAGTTTGAAGGTCTCTTTGCT 2374  
3294 GCGCTAGAGTCTATTCAAATTTTGGGTGTCTATTTCTTTATGGTGTTCATGACACT 3350  
2375 AAACGTGAGAGAGAGAGAGAGCTGGGATTTGCCAGCTTTGGGGCATCCATCACCACC 2434

3351 TTGGAAGACGTATT 3364  
2435 ATGGAGAGTCTT 2448  
RESULT 3  
US-08-762-500-74  
; Sequence 74, Application US/08762500  
; Patent No. 6030806  
; GENERAL INFORMATION:  
; APPLICANT: Landes, Gregory M.  
; APPLICANT: Burn, Timothy C.  
; APPLICANT: Connors, Timothy D.  
; APPLICANT: Dackowski, William R.  
; APPLICANT: Van Raay, Terence J.  
; APPLICANT: Klinger, Katherine W.  
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES, AND USING SAME  
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES, AND USING SAME  
; NUMBER OF SEQUENCES: 83  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENZYME CORPORATION  
; STREET: One Mountain Road  
; CITY: Framingham  
; STATE: Massachusetts  
; COUNTRY: United States of America  
; ZIP: 01701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/762,500  
; FILING DATE: 09-DEC-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/665,259  
; FILING DATE: 17-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/10469  
; FILING DATE: 17-JUN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dugan, Deborah A.  
; REGISTRATION NUMBER: 37,315  
; REFERENCE/DOCKET NUMBER: IGS-9.3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508) 872-8400  
; TELEFAX: (508) 872-5415  
; INFORMATION FOR SEQ ID NO: 74:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6525 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 573..5684  
; US-08-762-500-74  
Query Match 2.4%; Score 159.2; DB 3; Length 6525;  
Best Local Similarity 48.0%; Pred. No. 8.3e-27;  
Matches 554; Conservative 0; Mismatches 588; Indels 12; Gaps 3;  
2214 ATGTCACACTTAATAGTATATTTATGTCCTTCTGGCTGTCTATCTTGAATCAAGTCA 2273  
1935 ATGCTCTGCTGACTCTGCTCTATGCTGCTGACCTGGTACATGAGGCGGCTTC 1994  
2274 CAGGGGAATTTGGCTTACGAGATCATCTTTATATTTTCTGAAGCTTCATATTTGTC 2333  
1995 CAGGGCAGTTCGGCTGCTGCTCAGCCCTGGTACTTCTTCATCATGCCCTCTATTTG 2054  
2334 AAGAGTAAAGAAATTTATGAGGAGTTATCAGAGGCAATGTTAATGGAATATATAGTT 2393







APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION/DOCKET NUMBER: 36,373

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 264:

SEQUENCE CHARACTERISTICS:

LENGTH: 795 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-961-527-264

Query Match 1.2%; Score 80.6; DB 4; Length 795;

Best Local Similarity 48.3%; Pred. No. 3.6e-09;

Matches 320; Conservative 0; Mismatches 334; Indels 9; Gaps 3;

QY 4945 AGAAGTGGCACTAATAATCATCTCTTCTGTGTGAAAAAGAGAGATCTTAGGACTAT 5004

Db 3 ATAAAGAGGCGCTTGGAGTAATCTCAATTTGCAGATTGAAATGGAGAGATTATGGGCTTGA 62

QY 5005 TGGGTCCAAATGGTGGCGCAAGACAAATTAATTAATTTCTGTGGTGGTGAATTGAAC 5064

Db 63 TTGGTCATAATGGGCTGGGAAATTCGACCACTATAAATCCCTAGTCAGTATCATTTTCAC 122

QY 5065 CAACCTTCAGGCCAGGTATTTTATAGGAGATTATTTCTTCAGAGACAAGTGAAGATGATGATT 5124

Db 123 CCAGCAGTGTCTATTTTGGTAGACGGTCAGGAGTATTCGGAATCGCTGCTATTAA 182

QY 5125 CACTGAAGTGTAGGGTACTGTCTCAGATAAACCCCTTTGTGGCAGATACACTACATGTC 5184

Db 183 AACGAAAG---ATTGGCTACGTAGCAGACTCGCTGACTTATTTTACCGTTAACGCCCA 239

QY 5185 AGGAACAATTTTGAATTTATGGAGCTGTCAAAGGAATGAGTCAAGTGCATCAAGAGAAG 5244

Db 240 ATGAATTTTGGGAATTGATCGCCTCATCCATGATCTGAGTAGACTGCTGAGGAGCTA 299

QY 5245 TCATAAGTCGAATAACACATGCACCTTGATTTAAAAAGAACATCTTCAGAAGACTGTAAAGA 5304

Db 300 GTCTAGCTAGCTATTGAACGTTTTTGATTTTGTGTAATCGCTATCAGGTTATTGAAA 359

QY 5305 AACTACCTGAGGAATCAAAAGAAAGTTGTTTGTCTTAAGTAGTGTAGGGAATCCTC 5364

Db 360 CTCTTTCTCAGGAATCGTCAAGAAAGTCTTTGTCTATCGGAGCACTCTTGTCTGATCCCG 419

QY 5365 AGATTACTTTGTAGTAGTGAACCATCTACAGGTATGATGATCCCAAGCCAAACACACATGT 5424

Db 420 ATATTTGGGCTTGGATGAACCTTTGACTGGTTGGATCCCAAGGCTGCTTTGA---TT 476

QY 5425 GCGGAGCAATTCGAATTCGCAATTTAAAAACAGAAAGCGGGCTGCTATTTCTGACCACTCACT 5484

Db 477 TGAACAGATGATCAAGGAACATCCACAAAGGGAAGACAGTCTGTTTCAACTCATG 536

QY 5485 ATATGGAGGAGGAGAGGCTGTCTGTGATCAGTAGCTATCATGTTGTCTGGGCAAGTTAA 5544

Db 537 TCCTAGAGGTGGCAGACAAAGTCTGTGATCGGATGGCCATTTTGAAGAAAGGGGCATTGA 536

QY 5545 GATGTATCGGAACAGTACACATCTAAAGAGTAAATTTGAAAGGCTACTTTTTCGAA 5604

Db 597 TTTATTGTGTAGTGTAGAGGACTT---GAGAAAGATTACCCAGACCACTCTTTGAAA 653

QY 5605 TTA 5607

Db 654 GTA 656

RESULT 7

US-09-134-001C-987

; Sequence 987, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 987

; LENGTH: 696

; TYPE: DNA

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-987

Query Match 1.2%; Score 77.2; DB 4; Length 696;

Best Local Similarity 48.0%; Pred. No. 2e-08;

Matches 285; Conservative 0; Mismatches 303; Indels 6; Gaps 2;

QY 4936 GAAAAGTAAAGAAAGTGGCACTAAATACATCTCTTCTGTGTGAAAAAGAGAGATCT 4995

Db 50 GATATGGAATAAAACCTGTAATTTAAAGATTTGAATTTTGAATTAGAAAAAGCTGAAATTG 109

QY 4996 TAGGACTATTGGTCCAAATGGTGTGCGCAAGCACAATTAATTAATTTCTGTTGGTG 5055

Db 110 TAGGTTAATAGTCTAAATGGTGTGCGAAAGTACTACAATTAACATATGCTAGGAT 169

QY 5056 ATATTGAACCAACTTTCCAGCCAGGTATTTTATAGGAGATTATTTCTTCAGAGACAAGTGAAG 5115

Db 170 TGTATTAAATCCAAATGGAAGCAAAATTAATCAATTTCAATTAATTAATTAATGAATATTG 229

QY 5116 ATGATGATTCACTGAAGTGTATGGGTACTGTCTCAGATAAACCCCTTTGTGGCCAGATA 5175

Db 230 AAAATTATAGAAAGAAAGT---TATCTTATATACCTGGAATCACCAGTCACTATGATGAAC 286

QY 5176 CTACATTGCGAGCAATTTTGAATTTATGGAGCTGTCAAAGGAATGAGTCAAGTGACA 5235

Db 287 TAACTTTTAGAGGAGCACATTTGAAATGACGGCGATGGCTTATCAATTTAAGTAGAGAGAGG 346

QY 5236 TGAAGAAGTCTAATAGTCCGAATAACACATGCACTTTGATTTAAAGAAACATCTTCAGAAGA 5295

Db 347 TTATCGGTAGACTAAACCAATTAATTAAGTTTTTAGATTAGAAATGAACTTAAAGTAT 406

QY 5296 CTGTAAGAAGAACTACCTCGAGGAATCAAAAGAAAGTTGTGTTTGTCTTAAGTATGCTAG 5355

Db 407 TTCCAAGTCATTTTCTAAAGGGATGAAACAAAGAGTATGATGATTATTTGTGCAATTTATAG 466

QY 5356 GGAATCCTCAGATTACTTTGTAGTAGAACCATCTACAGGTATGNTCCCAAGCCCAAC 5415

Db 467 TTGATCCGAGATTATATATTTATTTGACAACTTTCTCGGTTTAGATCTTTGGGGATAC 526

QY 5416 AGCAGATGCGGAGCAATTCGAATTCGAATTTAAAAACAGAAAGCGGGCTGCTATTCTCA 5475

Db 527 AGTCTATGTAGATTATATGGA---GAAAAAGAAATGAAATCGAACAGATTATTAATGA 583

QY 5476 CCACTCAGTATATGAGGAGGAGGAGGCTGTCTGTGATGAGTAGCTATCATGCG 5529

Db 584 GTACACATATTTTGGCTACAGCCGAAAGGTATTTGTGATAGATTATTTATTATTAG 637

RESULT 8

US-09-134-001C-1205

; Sequence 1205, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1205
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1205

Query Match 1.2%; Score 75.8; DB 4; Length 1026;
Best Local Similarity 47.8%; Pred. No. 4.9e-08;
Matches 255; Conservative 0; Mismatches 272; Indels 6; Gaps 1;

QY 2456 TATTCAGAACACATACAGAAAGAGGTGAAATGTGGAGGCTTTGAGAAATTTGTCATT 2515
Db 18 TGTAAACAAAGTTTTTCGCAAAAAGAGAACTATTCAAGCTTTGAAAAATGTATCAAT 77
QY 2516 TGACATATATGAGGTCAGATTACTGCTTACTTTGGCCACAGTGGACAGAAAGAGTAC 2575
Db 78 TAGATTGACCAACATGATATTTTGGTGTATTGGTTATGTAGTGTGCTGGTAAAGTAC 137
QY 2576 ATTGATGAATATTTCTTGTGGACTCTGCCACCTTCTGTATGGGTTTGCATCTATATATGG 2635
Db 138 ATTAGTTCGGTTAGTCAATCACTTGAGACAGTATCAGATGTCAAGTTATTGTTGATGG 197
QY 2636 ACACAGATGCT-----TCGAAATAGATGAAATTTTGAAGCAAGAAATAGATTGGCAT 2689
Db 198 TCATGAGATTGATACATATAAAGAAAAAGATTTACGTGATATTAAAAAAAGATATCGGTAT 257
QY 2690 TTGTCCACAGTAGATATACATTTTGATGTTTTCACAGTAGAAGAAATTTATCAATTT 2749
Db 258 GATCTTTCAACATTTCAATTTGCTTAATTTCTAAATCAGTCTTATAAAATGTGCAATGCC 317
QY 2750 GGTCTTCAATCAAGGGGATACCGCAACAATAATAACAAGAGTGCGAAGGTTTACT 2809
Db 318 ACTTATTTTAAGTAAGCAATAAAGAAAGAAATTAAGAAAGTTGACGCAATGTTAGA 377
QY 2810 AGATTAGACATGACAGACTATCAAGATAACCAAGCTTAAATAATTAAGTGGTGTCAAAA 2869
Db 378 ATTGTGGGGCTTCTGATATAAAGATCAATTTCCAGATGAATTTATCAGTGGACAAAA 437
QY 2870 AAGAACTGTTCATGGAATGCTGTTCTTGGGAACCAAGATCTGCTGCTAGATGA 2929
Db 438 ACAACGTGTTGCCATCGCAAGAGCATTAGTAACGCATCCTTAAATATATTATTATGTGATGA 497
QY 2930 ACCAACAGCTGGAATGACCCCTGTTCTCGACATATTGTATGGAATCTTTTAA 2982
Db 498 AGCGACAAGTGTCTGATCCAGCTACTACAGCTCAATTTTAAATTTATTA 550

RESULT 9
US-09-107-532A-412
; Sequence 412, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 412:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 774 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...774
; SEQUENCE DESCRIPTION: SEQ ID NO: 412:
US-09-107-532A-412

Query Match 1.1%; Score 75; DB 4; Length 774;
Best Local Similarity 47.0%; Pred. No. 6.7e-08;
Matches 302; Conservative 0; Mismatches 335; Indels 6; Gaps 2;

QY 4934 AAGAAAGTAAAGAAAGTGCACATAAATACATCTCTTCTGTGTGAAAAAGGAGAT 4993
Db 57 AGTTAGGGAAAAAACAAGCATTAAGAAGCTTCTCTCTCGATAGACAGGGGAAT 116
QY 4994 CTTAGGACTATTGGGTCCAAATGGTCTGCGCAAAAGCAAAATATTAAATTTCTGGTTGG 5053
Db 117 TTTTGGATTTTAGGACATAATGTCGCGGAAATCAACTACGATTAAAGATTTAGTTAG 176
QY 5054 TGATATTGAACCAACTTCAGGCCAGGTATTTTAGAGATATTCTTCAGAGACAAGTGA 5113
Db 177 TATTATTGAACCATCATCGGAAACGATTTTAGTAGATGGAATGAAATTCGCGGAAATCG 236
QY 5114 AGATGATGATTCACTGAAGTGTATGGTTACTGTCTCAGATAAACCCCTTTTGTGGCAGA 5173
Db 237 ATTGAGTATCAAAACAAAA---AATCGGTATGTGCTGATTCGCCAGATATTTTTTACA 293
QY 5174 TACTACATTGCGAGAAACATTTTGAAATTTATGAGCTGTCAAAAGGAATGAGTCAAGTGA 5233
Db 294 ATTGACGGCTGGAGATATTGGGACTTGATTTTCAGCTGCTATGATTAAGACACACAAA 353
QY 5234 CATGAAAGAGTCAATAGTCGAATAACACATGCACTTGATTTAAAGAACATCTTCAGAA 5293
Db 354 AAAAGAAAAAGCGTTTAGCAGAAATGACAGCGCTTTTGTATATGATATGATCATCAAAATGA 413
QY 5294 GACTGTAAAGAAACTACCTGCAGGAATCAACGAAAGTTGTGTTTGTCTTAAGTATGCT 5353
Db 414 AACAAATCGCAAGTTTTTCACATGGAATCGGTCAAAAACCAATTTTGAATCGGCAGTTGCT 473
QY 5354 AGGGAATCCTCAGATTACTTTTGTAGTAGAACCATCTACAGGTATGATGCCAAAGCCAA 5413
Db 474 GCCAGATCCAGATATTGGGTATTGGATGAACTTTTGAAGGACTTGATCTCAGACAGC 533
QY 5414 ACAGCACATGTGGCGAGCAATTCGAATTTAAAGAACAGAAAGCGGCTGCTATCT 5473
Db 534 ATTTGATTGAAAGAAATGATGAAGGCTCATGCGCAAAAGGAAA---AACAGTCAATTT 590



LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349491)..(1349491)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1470091)..(1470091)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1603734)..(1603734)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1664854)..(1664854)  
OTHER INFORMATION: n equals a, t, c, or g  
US-08-916-421B-1

Query Match 1.1%; Score 74.6; DB 4; Length 1664976;  
Best Local Similarity 49.0%; Pred. No. 1.3e-06;

Matches 259; Conservative 0; Mismatches 264; Indels 6; Gaps 2;

QY 2493 GAGCTTTGAGAAATTTGTCATTTGACATATATAGGGTGCAGATTTACTGCTTACTTGGC 2552  
DB 444438 GCGGTTTTAAAGGAATAAATTTTAAAGTAAAGGAAGGAATGGTCTCTTTACTCGGC 444497  
QY 2553 CACAGTGGACAGAGAGATGACATTCATGATATATTTTGTGGACTCTGCCACCTTCT 2612  
DB 444498 CCTAATGGAGCTGGAAATCAACCTTTATTTTACCTCAATGGAATTTCTAAGACCTACA 444557  
QY 2613 GATGGGTTGCATCTATATATGG---ACACAGAGTCTCAGAAATAGATGAATGTTTGA 2669  
DB 444558 AAGGAGAGGTTTAAATAAGGCAAGCAATTAATATGATTAATAAGGCTTGGTGAA 444617  
QY 2670 GCAAGAAATAATGATGG---CATTTGTCCAGTTAGATATACATTTGATGTTTTGACA 2726  
DB 444618 GTTAGAAGACGGTTGATGGTTTTCAGAAATCCCGATGATCAGATATTTGCCCTTACA 444677  
QY 2727 GTAGAAGAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAATATATA 2786  
DB 444678 GTTAAGGAGGACGCTGGCAATTTGGACCTTTAAATCTTGGCTTGCCTAAAGGAAGTTGAG 444737  
QY 2787 CAAGAAGTGACAGAGTTTACTAGATTTAGACATGACAGACTATCAAGATAACCAAGCT 2846  
DB 444738 AAGAGGTTTAAGAGCGGTTTAAAGCTGTAGGAATGAAGGTTTGAATAAATCACTCCT 444797  
QY 2847 AAAAAATTAAGTGTGTGTCACAAAAGAAAGCTGTCAATTAGGAATTTGCTTCTTGGGAAC 2906

DB 444798 CATCATTTAAGTGGAGGACAAAAAAGAGAGTGGCTATATAGCAGTATTTTAGCTATGCAG 444857  
QY 2907 CCAAGATATCTGCTAGATGAACCAACAGCTGGAATGACCCCTGTTCTCGACATATT 2966  
DB 444858 CCTGAGGTTATGTTTTGGATGACCAACAGCTGGCTTAGACCTGTTGGAGCATCAAA 444917  
QY 2967 GATGAATCTTTTAAATACAGAAAGCAATCGGGTGACAGTGTTC 3015  
DB 444918 ATAATGAAATCTCTATACGATTTGAATAAAAGGCGATGACCAATAATA 444966

## RESULT 11

US-09-107-532A-1681  
Sequence 1681, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSER: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 1681:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 768 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...768  
SEQUENCE DESCRIPTION: SEQ ID NO: 1681:  
US-09-107-532A-1681

Query Match 1.1%; Score 72.2; DB 4; Length 768;  
Best Local Similarity 48.5%; Pred. No. 2.9e-07;  
Matches 237; Conservative 0; Mismatches 243; Indels 9; Gaps 1;

QY 2472 AGAAAGAGGTCAGAAATGAGGCTTTCAGAAATTTGTCATTTGACATATATAGGGT 2531  
DB 61 AAAAAAATCAGAAAAAGAGCGAAGCATTTAAGCGGCATCAGCTTCGATGTAGAAAAAGGC 120  
QY 2532 CAGATTACTGCTTACTTGGCCACAGTGGAAAGAGTACATTGATGATATCTTT 2591



```

; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664854)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1

```

```

Query Match 1.1%; Score 71; DB 4; Length 1664976;
Best Local Similarity 46.4%; Pred. No. 8.7e-06;
Matches 276; Conservative 0; Mismatches 310; Indels 9; Gaps 1;

Qy 2430 GGAAGAAGCCATAAGAAATAGTGGTATTCAGAGACATACAGAAAGAGGGTGAAAT 2489
Db 1656126 GGAATATTTTGTAGAGAGCTAAATGTATGGTAAAGGGAAGCAAA 1656067

```

```

Qy 2490 GTGGAGGCTTTGAGAAATTTGTCAATATATGAGGGTCAGATTACTGCTTACTT 2549
Db 1656066 ACCATAGCTTTAAAAATATTAATTTAAAAATTTGAAGAGGAGAGTTTGTAAATGATAATG 1656007
Qy 2550 GGCCACAGTGGACAGAAAGAGTCAATCATCAATATTTCTTTGTGGACTTCGCCACCT 2609
Db 1656006 GGGCCAGTGGATGTGGAAATCTAGCTTTATTTAAATATTTAGCTCTCTTAGATACTCCA 1655947
Qy 2610 TCTGATGGGTTTGCATCTATATATGACACAGAGTCTCGAAATAGATGAAATGTTTGA 2669
Db 1655946 ACAAGGGGAGGCTTTATATATAAGGGAGAGAAACAAGTTCAATGAGTGAATGAAGG 1655887
Qy 2670 GCAAGAAAAATGATGGCATTTGTCACAGTTAGATATACATTTGATGTTTTGACAGTA 2729
Db 1655886 GCTATATTTAGGAGAAAGATTAGTGATTTATATTTTCCAGCAATCCACTTAATAAAAAACA 1655827
Qy 2730 GAAGAAAAATTTCAATTTTGGCTTCAATCAAGAGGATACCAGCAACATATAATACAA 2789
Db 1655826 TTAAACAGCCTTAGAAAAATGTTGAATTTACCAATGATGTTAGATGAGAGATAAGATTAT 1655767
Qy 2790 GAAGTGCAGAAAGGTTTTTACTAGATTTAGACATGCAGACTATCAAGATA-----AC 2840
Db 1655766 AGAAGGAAGAGGGCAAAAAGCTTTTGGAGATGGTTGGTTTAGGAGATAGTTGAATCAT 1655707
Qy 2841 CAAGCTAAAAAATTAAGTGGTGTCAAAAAGAAAGCTGTCTATAGGAATGCTGTTCTT 2900
Db 1655706 TACCTTCATCAATTTGAGTGGAGGGCAACAACAAAGAGTTGCTATAGCAAGGGCTTTAGCA 1655647
Qy 2901 GGGAACCCAAAGATCTGCTAGATGAACCAACAGCTGGAATGACCCCTGTTCTCGA 2960
Db 1655646 AACACCCCAAAATATATTTGCTGATGAGCCAAACAGGAATTTGGACAGCAAAAGTGA 1655587
Qy 2961 CATATTGTATGGAATCTTTTAAATATACAGAAAGCCAAATCGGTGACAGTGTTC 3015
Db 1655586 ATGGCTGTATGAGTATCTTTAAAGGGTTGAATGAAAAGGAATACATAATTA 1655532

RESULT 13
US-09-107-532A-1190
; Sequence 1190, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277

```







**THIS PAGE BLANK (USPTO)**





Qy	1681	ACCTAGTTATAGCAATTTTCACTTTTGGATACCTTTTGGCAATTCATATCGTAGCAGAAA	1740
Db	1681	ACCTAGTTATAGCAATTTTCACTTTTGGATACCTTTTGGCAATTCATATCGTAGCAGAAA	1740
Qy	1741	AAGAAAAAATAAAGAAATTTTAAAGATAAATGGGACTTCATGATACTGCCTCTTTGGC	1800
Db	1741	AAGAAAAAATAAAGAAATTTTAAAGATAAATGGGACTTCATGATACTGCCTCTTTGGC	1800
Qy	1801	TTTCTGGGTCTCTATATACAAGTTAAATTTTCTATGTCCCTCTTATGGCAGTCA	1860
Db	1801	TTTCTGGGTCTCTATATACAAGTTAAATTTTCTATGTCCCTCTTATGGCAGTCA	1860
Qy	1861	TTGGCAGCTTCTTTGTTATTTTCTCAAAAGTAGCAGCATTTGATATTTCTGCTTTTTT	1920
Db	1861	TTGGCAGCTTCTTTGTTATTTTCTCAAAAGTAGCAGCATTTGATATTTCTGCTTTTTT	1920
Qy	1921	TCCTTTATGATATCATCTGTATTTTTCGTTTAAATGTCGACACCTCTTTTTTAAAAAAT	1980
Db	1921	TCCTTTATGATATCATCTGTATTTTTCGTTTAAATGTCGACACCTCTTTTTTAAAAAAT	1980
Qy	1981	CAAAAATGCGGAATAGTTGAATTTTGTACTGTGGCTTTTGGATTTATTTGGCCTTA	2040
Db	1981	CAAAAATGCGGAATAGTTGAATTTTGTACTGTGGCTTTTGGATTTATTTGGCCTTA	2040
Qy	2041	TGATATCCCTCATAGAAAGTTTCCCAATCGTTAGTGGCTTTTCAGTCCCTTTCTGTC	2100
Db	2041	TGATATCCCTCATAGAAAGTTTCCCAATCGTTAGTGGCTTTTCAGTCCCTTTCTGTC	2100
Qy	2101	ACTGTACTTTGTGATTTGTCACAGCTCATGATTTAGAAGATTTTAAATGAAGGTG	2160
Db	2101	ACTGTACTTTGTGATTTGTCACAGCTCATGATTTAGAAGATTTTAAATGAAGGTG	2160
Qy	2161	CTTCAATTTTCAAAATTTGACTGCGAGGCCATATCTCTAAATTAATACAAATCATGCTCA	2220
Db	2161	CTTCAATTTTCAAAATTTGACTGCGAGGCCATATCTCTAAATTAATACAAATCATGCTCA	2220
Qy	2221	CACTTAATAGTATATCTATGTCTCTGCTGCTCATCTTGATCAAGTCATTTCCAGGGG	2280
Db	2221	CACTTAATAGTATATCTATGTCTCTGCTGCTCATCTTGATCAAGTCATTTCCAGGGG	2280
Qy	2281	AAATTTGGCTTACGGAGATCATCTTTATATATTTCTGAAGCCTTCATATGGTCAAGAGTA	2340
Db	2281	AAATTTGGCTTACGGAGATCATCTTTATATATTTCTGAAGCCTTCATATGGTCAAGAGTA	2340
Qy	2341	AAAGAAATTAAGGAGTTATCAGAGGCCAATGTTAATGGAATAATAGTTTGTAGTGAAA	2400
Db	2341	AAAGAAATTAAGGAGTTATCAGAGGCCAATGTTAATGGAATAATAGTTTGTAGTGAAA	2400
Qy	2401	TTATTGAGCCAGTTTCTTTCAGAAATTTGTAGAAAAGAACCCATAGAATTTAGTGGTATTC	2460
Db	2401	TTATTGAGCCAGTTTCTTTCAGAAATTTGTAGAAAAGAACCCATAGAATTTAGTGGTATTC	2460
Qy	2461	AGAAGACATACAGAAAAGAGGGTGAAATGTGGAGGCTTTGAGAAAATTTGTCAATTTGACA	2520
Db	2461	AGAAGACATACAGAAAAGAGGGTGAAATGTGGAGGCTTTGAGAAAATTTGTCAATTTGACA	2520
Qy	2521	TATATAGGGTCAGATTACTGCTTACTTGGCCACAGTGGAAACAGGAAGAGTACATTGA	2580
Db	2521	TATATAGGGTCAGATTACTGCTTACTTGGCCACAGTGGAAACAGGAAGAGTACATTGA	2580
Qy	2581	TGAATATTTCTTTGAGACTCTGCCCACTTCTGATGGTTTTGTCATCTATATATGGACACA	2640
Db	2581	TGAATATTTCTTTGAGACTCTGCCCACTTCTGATGGTTTTGTCATCTATATATGGACACA	2640
Qy	2641	GAGTCTCAGAAAATAGATGAAATGTTTGAAGCAGAAAAAATGATTTGGCAATTTGTCCACAGT	2700
Db	2641	GAGTCTCAGAAAATAGATGAAATGTTTGAAGCAGAAAAAATGATTTGGCAATTTGTCCACAGT	2700
Qy	2701	TAGATATACACTTTGATGTTTTGACAGTAGAGAAATTTTATCAATTTTGGCTTCAATCA	2760
Db	2701	TAGATATACACTTTGATGTTTTGACAGTAGAGAAATTTTATCAATTTTGGCTTCAATCA	2760

Qy	2761	AAGGGATA	CCAGCCACA	ATAATA	TACAAGA	GTGCAGAA	GGTTTTT	TACTAGAT	TTT	TAGACA	2822							
Db	2761	AAGGGATA	CCAGCCACA	ATAATA	TACAAGA	GTGCAGAA	GGTTTTT	TACTAGAT	TTT	TAGACA	2820							
Qy	2821	TGCAGACT	ATCAAGAG	TAAACC	AAAGCT	TAAAAA	TTTAAG	TGGT	CGTCAAAA	AGAAAG	CTGT	2880						
Db	2821	TGCAGACT	ATCAAGAG	TAAACC	AAAGCT	TAAAAA	TTTAAG	TGGT	CGTCAAAA	AGAAAG	CTGT	2880						
Qy	2881	CATTAGGA	ATTGCTGT	TTTGGGA	ACCCAA	AGATACT	GTCTGT	AGATGA	AAACCA	ACAGCTG	3290							
Db	2881	CATTAGGA	ATTGCTGT	TTTGGGA	ACCCAA	AGATACT	GTCTGT	AGATGA	AAACCA	ACAGCTG	3290							
Qy	2941	GAATGGA	CCCTTGT	TTCTG	CAATAT	TATG	TATG	TATG	TATG	TATG	3000							
Db	2941	GAATGGA	CCCTTGT	TTCTG	CAATAT	TATG	TATG	TATG	TATG	TATG	3000							
Qy	3001	GGGTG	CAGTGT	TTCAGT	ACTCAT	TTCAT	TGAT	TGNA	TCTTTT	TAAATAC	CAGAAA	AGCCAA	TC	3060				
Db	3001	GGGTG	CAGTGT	TTCAGT	ACTCAT	TTCAT	TGAT	TGNA	TCTTTT	TAAATAC	CAGAAA	AGCCAA	TC	3060				
Qy	3061	CTGTG	TATCAC	AAAGAA	TGCTG	AAAT	TGTTT	TGGT	TCTT	CAAT	TGTTT	CCTCA	AAAGT	TAAT	3120			
Db	3061	CTGTG	TATCAC	AAAGAA	TGCTG	AAAT	TGTTT	TGGT	TCTT	CAAT	TGTTT	CCTCA	AAAGT	TAAT	3120			
Qy	3121	GGGGAT	TCGGCT	ACCGCT	GTAG	CAAT	GATAG	CAAA	TAT	TGT	GC	CA	CAGAT	CTCT	TT	3180		
Db	3121	GGGGAT	TCGGCT	ACCGCT	GTAG	CAAT	GATAG	CAAA	TAT	TGT	GC	CA	CAGAT	CTCT	TT	3180		
Qy	3181	CTTCACT	TGGTTAA	CAAC	ATATAC	CTTGG	AGCTACT	TTTAT	TACA	CA	AGA	TG	ACCA	CA	AC	3240		
Db	3181	CTTCACT	TGGTTAA	CAAC	ATATAC	CTTGG	AGCTACT	TTTAT	TACA	CA	AGA	TG	ACCA	CA	AC	3240		
Qy	3241	TTGTG	TATAG	CTTGG	CTTT	CAAG	CAATGG	CAAAA	TTTTT	CAG	GTG	TTG	TTT	CT	CG	CCCT	TAG	3300
Db	3241	TTGTG	TATAG	CTTGG	CTTT	CAAG	CAATGG	CAAAA	TTTTT	CAG	GTG	TTG	TTT	CT	CG	CCCT	TAG	3300
Qy	3301	ACAGT	CAAT	CAAA	TTTTT	GGG	TGTCAT	TTT	TAT	TGG	TGTT	CC	ATG	AC	AG	CA	AG	3360
Db	3301	ACAGT	CAAT	CAAA	TTTTT	GGG	TGTCAT	TTT	TAT	TGG	TGTT	CC	ATG	AC	AG	CA	AG	3360
Qy	3361	TATTTT	TAAAG	CTAG	AAAGT	TGA	AGCA	AAAT	TG	CA	CA	AG	CA	AG	CA	AG	CA	3420
Db	3361	TATTTT	TAAAG	CTAG	AAAGT	TGA	AGCA	AAAT	TG	CA	CA	AG	CA	AG	CA	AG	CA	3420
Qy	3421	AGCAG	CCACT	TGG	AGGA	AAAT	TG	CAAT	CTTT	TG	AT	G	AAAT	TG	GA	AA	CA	3480
Db	3421	AGCAG	CCACT	TGG	AGGA	AAAT	TG	CAAT	CTTT	TG	AT	G	AAAT	TG	GA	AA	CA	3480
Qy	3481	TTAT	CTTT	CTG	AAAC	CA	AGG	CTT	CT	TAG	TG	AG	CA	CC	ATG	AG	CC	3540
Db	3481	TTAT	CTTT	CTG	AAAC	CA	AGG	CTT	CT	TAG	TG	AG	CA	CC	ATG	AG	CC	3540
Qy	3541	TGT	ATAC	ATAG	CA	AA	AGTTT	CA	TTT	CTT	TAC	CTT	TG	AA	AG	CTG	AA	3600
Db	3541	TGT	ATAC	ATAG	CA	AA	AGTTT	CA	TTT	CTT	TAC	CTT	TG	AA	AG	CTG	AA	3600
Qy	3601	CAG	TGTT	GTCT	CTG	CTTTT	TAA	TTTTT	T	CA	CGT	TC	CAG	AT	TTTT	T	T	3660
Db	3601	CAG	TGTT	GTCT	CTG	CTTTT	TAA	TTTTT	T	CA	CGT	TC	CAG	AT	TTTT	T	T	3660
Qy	3661	ACT	CTTT	TAAAA	TG	CTG	TTGG	TTC	CC	AT	CA	AA	CTG	TT	CC	AG	CTT	3720
Db	3661	ACT	CTTT	TAAAA	TG	CTG	TTGG	TTC	CC	AT	CA	AA	CTG	TT	CC	AG	CTT	3720
Qy	3721	CTG	GAG	CA	AA	CC	AT	TA	A	T	A	A	AA	AG	CT	CT	CTT	3780
Db	3721	CTG	GAG	CA	AA	CC												

|||||  
3841 ACAGTGACTATGATCGTGGCTCCCATAGTGGCTTTAAATGTGATGCAATTCAGAAA 3900  
QY  
3901 AGAGCTATGTTTTGGAGCTGTTTTCAAGTACTATGTTTTTAATCTTTACCTATATTAG 3960  
Db  
3901 AGGACTATGTTTTGGAGCTGTTTTCAACAGTACTATGTTTTTAATCTTTACCTATATTAG 3960  
QY  
3961 TGAATATCATTAGTAACTACTATCTTTATCATTTAAATGTGACTGAAACCATCCAGATCT 4020  
Db  
3961 TGAATATCATTAGTAACTACTATCTTTATCATTTAAATGTGACTGAAACCATCCAGATCT 4020  
QY  
4021 GGAGTACCCCATCTTTCAAGAAATTAAGTATAGTTTTTAAATTTGAGCTGATTTTTTC 4080  
Db  
4021 GGAGTACCCCATCTTTCAAGAAATTAAGTATAGTTTTTAAATTTGAGCTGATTTTTTC 4080  
QY  
4081 AAGCAGCTTTGCTGGAAATCAATTTGTTACTGCAATGACCATCTTACTTTGGCCATGGAAATG 4140  
Db  
4081 AAGCAGCTTTGCTGGAAATCAATTTGTTACTGCAATGACCATCTTACTTTGGCCATGGAAATG 4140  
QY  
4141 CAGAGATCATAGATCAAGCTTATGTTGATATCCCTTTATTTTATCACTTTTATTTTGA 4200  
Db  
4141 CAGAGATCATAGATCAAGCTTATGTTGATATCCCTTTATTTTATCACTTTTATTTTGA 4200  
QY  
4201 CATATTTGGATTTGCAAGCTGTTGTTGATATCCCTTTATTTTATCACTTTTATTTTGA 4260  
Db  
4201 CATATTTGGATTTGCAAGCTGTTGTTGATATCCCTTTATTTTATCACTTTTATTTTGA 4260  
QY  
4261 TGCTAGGAAGCTTACTGCGCATTTCAATATGGAATATATTTTTATATGTAAGTTCCTTG 4320  
Db  
4261 TGCTAGGAAGCTTACTGCGCATTTCAATATGGAATATATTTTTATATGTAAGTTCCTTG 4320  
QY  
4321 CTGTGGTTTTTGGCTTTATTTGTTGATGTTCCATCAGTTATTTCTGTCACATATATTGCTT 4380  
Db  
4321 CTGTGGTTTTTGGCTTTATTTGTTGATGTTCCATCAGTTATTTCTGTCACATATATTGCTT 4380  
QY  
4381 CTTTCACCTTTAAGAAATTTTAAATCAAGAAATTTTGGTCATTTATCTATTTCTGTGG 4440  
Db  
4381 CTTTCACCTTTAAGAAATTTTAAATCAAGAAATTTTGGTCATTTATCTATTTCTGTGG 4440  
QY  
4441 CAGCGTTGNCCTGATTTGCAATCACTGAAATATCTTTCTTTATGGGATACACAATTCGAA 4500  
Db  
4441 CAGCGTTGNCCTGATTTGCAATCACTGAAATATCTTTCTTTATGGGATACACAATTCGAA 4500  
QY  
4501 CTATTTCTTCAATATGCTTTGTTGATCATCTTCAATCTACCTTCTAGGTTGCTGA 4560  
Db  
4501 CTATTTCTTCAATATGCTTTGTTGATCATCTTCAATCTACCTTCTAGGTTGCTGA 4560  
QY  
4561 TTTCTTTTCAATAAGATTTCTTGAAGAAATGTAAGAAATTTGGAACCATCTATAATCCAT 4620  
Db  
4561 TTTCTTTTCAATAAGATTTCTTGAAGAAATGTAAGAAATTTGGAACCATCTATAATCCAT 4620  
QY  
4621 GGGATAGGCTTTCAATAGCTGTTATATCGCTTACCTGAGTGTGATCTGTGGATTTTCC 4680  
Db  
4621 GGGATAGGCTTTCAATAGCTGTTATATCGCTTACCTGAGTGTGATCTGTGGATTTTCC 4680  
QY  
4681 TCTTACATATGATGAGAAATATGAGCGAGATCAATAAGAAAGATCCCTTTTCA 4740  
Db  
4681 TCTTACATATGATGAGAAATATGAGCGAGATCAATAAGAAAGATCCCTTTTCA 4740  
QY  
4741 GAAACCTTTTCAAGAAATTTTAAATAGGAGCTTCCAGAAACCATCCAGACATGAGGATG 4800  
Db  
4741 GAAACCTTTTCAAGAAATTTTAAATAGGAGCTTCCAGAAACCATCCAGACATGAGGATG 4800  
QY  
4801 AAGATGAAGTGTCAAGCTGAAAGCTAAAGTCAAGAGCTGATGGGTTGCGAGTTT 4860  
Db  
4801 AAGATGAAGTGTCAAGCTGAAAGCTAAAGTCAAGAGCTGATGGGTTGCGAGTTT 4860  
QY  
4861 GTGAGGAGAACCATCTATGTCAGCAATTTGCAATGCAATGATGATGACAGAAAG 4920  
Db  
4861 GTGAGGAGAACCATCTATGTCAGCAATTTGCAATGCAATGATGATGACAGAAAG 4920  
QY  
4921 ATTTTCTTTTCAAGAAATGTAAGAAAGTGGCAACTAAATACATCTCTTTCTGTGTGA 4980  
Db

4921 ATTTTCTTTCTTCAAGAAAGTAAAGAAAGTGGCAACTAAATACATCTCTTTCTGTGTGA 4980  
QY  
4981 AAAAAAGGAGAGATCTTAGGACTATTGGCTCCAAATGGTGGCAAAAGACCAATATTA 5040  
Db  
4981 AAAAAAGGAGAGATCTTAGGACTATTGGCTCCAAATGGTGGCAAAAGACCAATATTA 5040  
QY  
5041 ATATTCTGTTGGTGATATTGAACCAACTTCAGGCCAGGTATTTTATGAGAGATTTCTT 5100  
Db  
5041 ATATTCTGTTGGTGATATTGAACCAACTTCAGGCCAGGTATTTTATGAGAGATTTCTT 5100  
QY  
5101 CAGAGCAAGTGAAGATGATGATTCACCTGAAGTGTATGGTTACTGTCTCAGATAAACC 5160  
Db  
5101 CAGAGCAAGTGAAGATGATGATTCACCTGAAGTGTATGGTTACTGTCTCAGATAAACC 5160  
QY  
5161 CTTTGTGCGCAGATACTACTATTGAGGAACTATTTTGAATTTATGAGCTGTCAAGGAA 5220  
Db  
5161 CTTTGTGCGCAGATACTACTATTGAGGAACTATTTTGAATTTATGAGCTGTCAAGGAA 5220  
QY  
5221 TGAGTGCAAGTGACATGAAAGAGTTCATAAGTTCGAATTAACACATGCACCTTGATTTAAAG 5280  
Db  
5221 TGAGTGCAAGTGACATGAAAGAGTTCATAAGTTCGAATTAACACATGCACCTTGATTTAAAG 5280  
QY  
5281 AACATCTTCAGAACTGTAAAGAACTACCTGCGAGGAACTCAAAACGAAAGTTGTGTTG 5340  
Db  
5281 AACATCTTCAGAACTGTAAAGAACTACCTGCGAGGAACTCAAAACGAAAGTTGTGTTG 5340  
QY  
5341 CTCTAAGTATGCTAGGGAATCCTCAGATTACTTTGCTAGATGAACCATCTCAGGTATGG 5400  
Db  
5341 CTCTAAGTATGCTAGGGAATCCTCAGATTACTTTGCTAGATGAACCATCTCAGGTATGG 5400  
QY  
5401 ATCCCAAGCCAAACAGACACATGTGGCGAGCAATTCGAATTCGCAATTTAAAAACAGAAAGC 5460  
Db  
5401 ATCCCAAGCCAAACAGACACATGTGGCGAGCAATTCGAATTCGCAATTTAAAAACAGAAAGC 5460  
QY  
5461 GGGCTGCTATTTCTGACCACTCATATATGAGGAGGAGAGGCTGTCTGTGATCAGTAG 5520  
Db  
5461 GGGCTGCTATTTCTGACCACTCATATATGAGGAGGAGAGGCTGTCTGTGATCAGTAG 5520  
QY  
5521 CTATCATGTGTCTGGGAGTTAAGATGTATCGGAACAGTACCAATCTTAAAGAGTAAAT 5580  
Db  
5521 CTATCATGTGTCTGGGAGTTAAGATGTATCGGAACAGTACCAATCTTAAAGAGTAAAT 5580  
QY  
5581 TTGGAAGAGGCTTACCTTTTGGAAATTTAAATTTGAAGACTGGATAGAAAACCTTAGAAGTAG 5640  
Db  
5581 TTGGAAGAGGCTTACCTTTTGGAAATTTAAATTTGAAGACTGGATAGAAAACCTTAGAAGTAG 5640  
QY  
5641 ACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTGAGGAAAGTTTTT 5700  
Db  
5641 ACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTGAGGAAAGTTTTT 5700  
QY  
5701 CTTCTATTTTGGCTTATAAATTTCCCTAAGGAGAGATGTTTCAGTCCCTTTTACAAATCTTTT 5760  
Db  
5701 CTTCTATTTTGGCTTATAAATTTCCCTAAGGAGAGATGTTTCAGTCCCTTTTACAAATCTTTT 5760  
QY  
5761 TTAAGCTGGAAGAACTAAACATGCTTTTGCATTTGAAGAAATATAGCTTTTCTCAAGCAA 5820  
Db  
5761 TTAAGCTGGAAGAACTAAACATGCTTTTGCATTTGAAGAAATATAGCTTTTCTCAAGCAA 5820  
QY  
5821 CATTTGGAACAGGTTTTGTAGAACTCCTCAATTAAGAACTAAGAGAGAGATATAGTTG 5880  
Db  
5821 CATTTGGAACAGGTTTTGTAGAACTCCTCAATTAAGAACTAAGAGAGAGATATAGTTG 5880  
QY  
5881 GAACTTTAAACAGCACATTTTGGTGGGAAACCAACAGAGATAGTAGTATTTTGA 5940  
Db  
5881 GAACTTTAAACAGCACATTTTGGTGGGAAACCAACAGAGATAGTAGTATTTTGA 5940  
QY  
5941 TTTTGTATTTGCTGCTTACTGCGGACTTCTTTTCTTTTCACTTAATTTTAACTTTGG 6000  
Db  
5941 TTTTGTATTTGCTGCTTACTGCGGACTTCTTTTCTTTTCACTTAATTTTAACTTTGG 6000  
QY  
6001 TTTTAAAAAGTTTTTATTTGGAATGTTAACTGAGAAACCAAGAACGCACTTGGAATTTTTC 6060  
Db  
6001 TTTTAAAAAGTTTTTATTTGGAATGTTAACTGAGAAACCAAGAACGCACTTGGAATTTTTC 6060

QY	6061	TAAGCTCCTTAATGAATGCTGTGGTGTGTGTTTCTTTTAAATAAAGCGTAT	6120
Db	6061	TAAGCTCCTTAATGAATGCTGTGGTGTGTGTTTCTTTTAAATAAAGCGTAT	6120
QY	6121	GTATATTAAGTGAAGCTGATGTTTGTATTAAGTATATTAAGTATATTAAGTATG	6180
Db	6121	GTATATTAAGTGAAGCTGATGTTTGTATTAAGTATATTAAGTATATTAAGTATG	6180
QY	6181	TCATCTTTTTCACCAATTCAGAAACAGTCTTCTGAAATTTGTGATTTAAAGCAATTTGAAT	6240
Db	6181	TCATCTTTTTCACCAATTCAGAAACAGTCTTCTGAAATTTGTGATTTAAAGCAATTTGAAT	6240
QY	6241	AGAATAGTATTTTAAAGTTATCTTTAAAGTTTATGCCATCTCTTTAAATAAGTACGTA	6300
Db	6241	AGAATAGTATTTTAAAGTTATCTTTAAAGTTTATGCCATCTCTTTAAATAAGTACGTA	6300
QY	6301	ATGTTCCAACTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAGC	6360
Db	6301	ATGTTCCAACTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAGC	6360
QY	6361	AATGTGAAGTTTCTGCTTCTCTCTTTTAAATTTCTAAATAAAGCCACTTTGAATGGAAT	6420
Db	6361	ANGTGAAGTTTCTGCTTCTCTCTTTTAAATTTCTAAATAAAGCCACTTTGAATGGAAT	6420
QY	6421	TGTCATCCGTAAGAGCTGAAGTGTGAAGCACTAGCAATCTCAATATAGAGATTGGAGAA	6480
Db	6421	TGTCATCCGTAAGAGCTGAAGTGTGAAGCACTAGCAATCTCAATATAGAGATTGGAGAA	6480
QY	6481	AGTTATATCCACTAGTGGCAGTCATTCATGATCAATTAAGTGAAT	6525
Db	6481	AGTTATATCCACTAGTGGCAGTCATTCATGATCAATTAAGTGAAT	6525
RESULT 2			
AAD33648			
ID	AAD33648 standard; cDNA; 6369 BP.		
AC	AAD33648;		
XX			
DT	01-JUL-2002 (first entry)		
XX	Human TRICH-3 cDNA.		
DE			
KW	Human; transporter and ion channel; TRICH-3; transport disorder; angina;		
KW	amyotrophic lateral sclerosis; cystic fibrosis; neuromuscular disorder;		
KW	cardiac disorder; polymyositis; diabetes; neurological disorder; cancer;		
KW	depression; schizophrenia; anaemia; Wilson's disease; Cushing's disease;		
KW	cell proliferated disorder; infertility; arteriosclerosis; gene therapy;		
KW	Alzheimer's disease; Parkinson's disease; Huntington's disease; allergy;		
KW	myasthenia gravis; multiple sclerosis; metabolic disorder; hypertension;		
KW	acquired immune deficiency syndrome; immunological disorder; scleroderma;		
KW	endocrine disorder; autoimmune thyroiditis; rheumatoid arthritis; goitre;		
KW	cardiac myopathy; amnesia; toxic myopathy; Addison's disease; infection;		
KW	epilepsy; mental disorder; myocarditis; Crohn's disease; Grave's disease;		
KW	muscle disorder; stroke; dementia; anxiety; AIDS; asthma; cirrhosis;		
gene; ss.			
XX			
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	1238..6166	
FT	/*tag= a		
FT	/product= "Human TRICH-3 protein"		
FT			
PN	WO200212340-A2.		
XX			
PD	14-FEB-2002.		
XX			
XX	01-AUG-2001; 2001WO-US24217.		
XX			
PR	03-AUG-2000; 2000US-223269P.		
PR	10-AUG-2000; 2000US-224456P.		

PR	18-AUG-2000; 2000US-226410P.		
PR	25-AUG-2000; 2000US-228140P.		
PR	31-AUG-2000; 2000US-230067P.		
PR	08-SEP-2000; 2000US-231434P.		
XX	(INCY-) INCYTE GENOMICS INC.		
PA	Yue H, Thornton M, Rankumar J, Tang YT, Azimzai Y, Baughn MR;		
XX	Yang J, Yao MG, Lal P, Wallia NK, Gandhi AR, Hafalia AJA;		
PI	Nguyen DB, Patterson C, Elliott VS, Tribouley CM, Lu DAM, Xu Y;		
PI	Reddy R, Hernandez R, Borowsky ML, Lo TP, Lu Y, Policky JL;		
PI	Greene BD, Sanjanwala MS, Raumann BE, Burford N, Ison CH, Lee EA;		
PI	Ding L, Das D, Kallick DA, Khan FA, Seilhamer JU;		
XX	WPI; 2002-206330/26.		
DR	P-PSDB; AAE21159.		
XX	New human transporters and ion channels polypeptides and		
PT	polynucleotides for diagnosing, preventing or treating transport,		
PT	neurological, muscle, immunological and cell proliferative disorders		
XX	Claim 77; Page 203-205; 230pp; English.		
PS	The invention relates to human transporter and ion channel polypeptides		
CC	designated TRICH and nucleic acid molecules encoding such polypeptides.		
CC	TRICH sequences are useful for diagnosis, treatment and prevention of		
CC	transport, muscle, neurological, immunological and cell proliferative		
CC	disorders. Transport disorders include akinesia, amyotrophic lateral		
CC	sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular		
CC	dystrophy, diabetes mellitus, diabetes insipidus, myasthenia gravis,		
CC	myocarditis, prostate cancer, cardiac disorders associated with		
CC	transport e.g. polymyositis, bradyarrhythmia, dermatomyositis, angina,		
CC	neurological disorders associated with transport e.g. amnesia, bipolar		
CC	disorder, depression, Tourette's disorder, schizophrenia, other disorders		
CC	associated with transport e.g. neurofibromatosis, sickle cell anaemia,		
CC	Wilson's disease, cataracts, infertility, hyperglycaemia, hypoglycaemia,		
CC	goitre, Cushing's disease, hypercholesterolaemia and cystinuria. Cell		
CC	proliferated disorders include cancer, actinic keratosis, cirrhosis,		
CC	arteriosclerosis, atherosclerosis, buritis, hepatitis and psoriasis.		
CC	Neurological disorders include Alzheimer's, Pick's and Parkinson's		
CC	disease, amyotrophic lateral sclerosis, epilepsy, stroke, Huntington's		
CC	disease, multiple sclerosis, dementia and other extrapyramidal disorder,		
CC	motor neuron disorder, prion disease, metabolic disease of the nervous		
CC	system and other developmental disorders of the central nervous system,		
CC	neuromuscular disorders, metabolic, endocrine and toxic myopathies,		
CC	periodic paralysis, mental disorders including mood, anxiety, and		
CC	immunological disorders include acquired immune deficiency syndrome		
CC	(AIDS), adult respiratory distress syndrome, Addison's disease,		
CC	allergies, asthma, atherosclerosis, osteoporosis, autoimmune haemolytic		
CC	anaemia, autoimmune thyroiditis, Crohn's disease, atopic dermatitis,		
CC	Grave's disease, glomerulonephritis, rheumatoid arthritis, scleroderma,		
CC	systemic lupus erythematosus, systemic sclerosis, ulcerative colitis,		
CC	haemodialysis, uveitis; viral, bacterial, fungal, parasitic, protozoal,		
CC	helminthic infections and trauma; and muscle disorders include cardiac		
CC	myopathy, myocarditis, polymyositis, arrhythmias and hypertension. The		
CC	TRICH polynucleotides are used in gene therapy. The present sequence		
CC	is human TRICH-3 cDNA.		
XX			
SQ	Sequence 6369 BP; 1893 A; 1112 C; 1246 G; 2118 T; 0 other;		
Query Match 93.8%; Score 6122.8; DB 24; Length 6369;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 6127; Conservative 0; Mismatches 8; Indels 0; Gaps 0;			
QY	1	AAATGTTGATATTTCTCTTAGCAGCTGTCAACAGGTTAGTTTCAGTCTAAGTTT	60
Db	228	AAATGTTGATATTTCTCTTAGCAGCTGTCAACAGGTTAGTTTCAGTCTAAGTTT	287
QY	61	CTACCCACATTTCTTGAATCTGATTTGTCTATTTTATTTTCAAAAACCTTTTGCAG	120
Db	288	CTACCCACATTTCTTGAATCTGATTTGTCTATTTTATTTTCAAAAACCTTTTGCAG	347
QY	121	TACCTTTTGTGTGTTTGTGTGTCCTTGCAGTGAACAGTCTGGATTGGACAGTGGT	180



Db 348 TACCTTTTGGTCTGTCTGTGTGCTTGCAGTGAACAGTCTGGATTGGACAGTGT 407  
Qy 181 CTGTCGTGTTAGTTCAGTTTCTCAAGCCCTTGTGTCACTAATAGGATTTGATGATG 240  
Db 408 CTGTCGTGTTAGTTCAGTTTCTCAAGCCCTTGTGTCACTAATAGGATTTGATGATG 467  
Qy 241 TCCAGCTTGGGAATTATTACAGGAATTAAACAACTTTTACAGTGCCTTCCCTGAGCTC 300  
Db 468 TCCAGCTTGGGAATTATTACAGGAATTAAACAACTTTTACAGTGCCTTCCCTGAGCTC 527  
Qy 301 TCTTTCTAATTTGTTCCCTCTTCTACTTTTGTCTTCCCTGTGGCTGTCTTCTATCCTCC 360  
Db 528 TCTTTCTAATTTGTTCCCTCTTCTACTTTTGTCTTCCCTGTGGCTGTCTTCTATCCTCC 587  
Qy 361 AGCCAGAGAGTAGTGTATTTTCTCCATGTGTGTACACATTTGTGTCAGCTGCAACAC 420  
Db 588 AGCCAGAGAGTAGTGTATTTTCTCCATTTGTGTACACATTTGTGTCAGCTGCAACAC 647  
Qy 421 CATATCCAGGCCCAATGTTAGGAGGTAGAGAAAGCAAAAGGATTTGGCTCATCC 480  
Db 648 CATATCCAGGCCCAATGTTAGGAGGTAGAGAAAGCAAAAGGATTTGGCTCATCC 707  
Qy 481 TCTTACAAAGATAGTTCATTTGAATAGAGAAAGGTTTCTGCTCCCTCAGAGTGTGGCT 540  
Db 708 TCTTACAAAGATAGTTCATTTGAATAGAGAAAGGTTTCTGCTCCCTCAGAGTGTGGCT 767  
Qy 541 GCACTAGGCTTTGTTACTGTAGTCTGGCCCTGTTACCATGGGATTTGCTGCAATGTGGGG 600  
Db 768 GCACTAGGCTTTGTTACTGTAGTCTGGCCCTGTTACCATGGGATTTGCTGCAATGTGGGG 827  
Qy 601 ATACAGAGAAATTCAGAAAGAAAGAAATTTGCTAATTTCTACATTTCCCTCCGAGCAAT 660  
Db 828 ATACAGAGAAATTCAGAAAGAAAGAAAGAAATTTGCTAATTTCTACATTTCCCTCCGAGCAAT 887  
Qy 661 AAGACCTTCCCTTGCCCATTTCTCAATTTCAAGGCTTAAGGCTTCTTCTGGAGCTGCCCTCTGT 720  
Db 888 AAGACTTCCCTTGCCCATTTCTCAATTTCAAGGCTTCTTCTGGAGCTGCCCTCTGT 947  
Qy 721 GGGCGGTTGGGAGATACCAAGAGAGAAAGTACCACTGTTGATATGTTGGTATTTCAA 780  
Db 948 GGGCGGTTGGGAGATACCAAGAGAGAAAGTACCACTGTTGATATGTTGGTATTTCAA 1007  
Qy 781 ATTCTGCTCTACCTATTTCACATGCTCTGTTTACTTTTTCAGAGCTGACAGATTTGCTGCT 840  
Db 1008 ATTCTGCTCTACCTATTTCACATGCTCTGTTTACTTTTTCAGAGCTGACAGATTTGCTGCT 1067  
Qy 841 CCATGCAATCTGTCCAGTTTCTTAAGAGAGACAGCTTTGGAGTATGCTTTAATCCATCTTAC 900  
Db 1068 CCATGCAATCTGTCCAGTTTCTTAAGAGAGACAGCTTTGGAGTATGCTTTAATCCATCTTAC 1127  
Qy 901 CTGGGACTGAACAGCTGCTTATTTTCCGTTAAAAATTAATCATGCACTTTACTGCGTGGC 960  
Db 1128 CTGGGACTGAACAGCTGCTTATTTTCCGTTAAAAATTAATCATGCACTTTACTGCGTGGC 1187  
Qy 961 TCCGGCTTTGTTGTTTGTCTTAAATAGTGTATTCAGAAACATGTCACATG 1020  
Db 1188 TCCGGCTTTGTTGTTTGTCTTAAATAGTGTATTCAGAAACATGTCACATG 1247  
Qy 1021 CAATTAGGGAGTATGAGTGTGGAGACAGACCAACACTTCTACTGAAGAAATTTACTTAA 1080  
Db 1248 CAATTAGGGAGTATGAGTGTGGAGACAGACCAACACTTCTACTGAAGAAATTTACTTAA 1307  
Qy 1081 TTAATGAGAAACCAAAAGAGTATGTTTCAGGAAATTTCTTTTCCACTATTTTAT 1140  
Db 1308 TTAATGAGAAACCAAAAGAGTATGTTTCAGGAAATTTCTTTTCCACTATTTTAT 1367  
Qy 1141 TTTTGGTTAATTAATTTAGCATATGTCATCCAAATAGAAATATGAGAGATGCTTAATA 1200  
Db 1368 TTTTGGTTAATTAATTTAGCATATGTCATCCAAATAGAAATATGAGAGATGCTTAATA 1427  
Qy 1201 TAGAACTCAATCTTATGGACAAGTTTACTCTTTCTAATCTTAATTTCTGGATATACTCCAG 1260

Db 1428 TAGAACTCAATCTTATGGACAAGTTTACTCTTTCTAATCTAATTTCTTGGATATATCTCCAG 1487  
Qy 1261 TGACTAATATTACAGCAGCATCATGCAAGAGTGTCTACTGATCATCTACTACTGATGCA 1320  
Db 1488 TGACTAATATTACAGCAGCATCATGCAAGAGTGTCTACTGATCATCTACTGATGCA 1547  
Qy 1321 TAAATTACTGAAGAAATATACAAATGAAAAAGAAATGTTTAAACATCCAGTCTCTCTAAAGCCGA 1380  
Db 1548 TAAATTACTGAAGAAATATACAAATGAAAAAGAAATGTTTAAACATCCAGTCTCTCTAAAGCCGA 1607  
Qy 1381 GCAACTTTGTAGGTGTGTTTCAAGACTCCATGCTCTATGAACTTCGTTTCTTCTTCTG 1440  
Db 1608 GCAACTTTGTAGGTGTGTTTCAAGACTCCATGCTCTATGAACTTCGTTTCTTCTTCTG 1667  
Qy 1441 ATATGATTTCCAGTATCTTCTATTTATATGGAATTCAGAGCTGGCTGTTCAAAATCATGTG 1500  
Db 1668 ATATGATTTCCAGTATCTTCTATTTATATGGAATTCAGAGCTGGCTGTTCAAAATCATGTG 1727  
Qy 1501 AGGCTGCTCAGTACTGTGCTCCTCAGGTTTTCACAGTTTTCACAGCATCCATAGATGCTGCCA 1560  
Db 1728 AGGCTGCTCAGTACTGTGCTCCTCAGGTTTTCACAGTTTTCACAGCATCCATAGATGCTGCCA 1787  
Qy 1561 TTATACAGTTGAAGACCAATGTTTCTTTTGAAGGAGCTGGAGTCAACTAAAGCTGTTA 1620  
Db 1788 TTATACAGTTGAAGACCAATGTTTCTTTTGAAGGAGCTGGAGTCAACTAAAGCTGTTA 1847  
Qy 1621 TTATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTAAATTTAAATAT 1680  
Db 1848 TTATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTAAATTTAAATAT 1907  
Qy 1681 ACCTAGTTTATAGCATTTTCCACCTTTTGGATATCTTTTGGCAATTCATATCGTAGCAGAAA 1740  
Db 1908 ACCTAGTTTATAGCATTTTCCACCTTTTGGATATCTTTTGGCAATTCATATCGTAGCAGAAA 1967  
Qy 1741 AAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATFACTGCTTTTGGC 1800  
Db 1968 AAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATFACTGCTTTTGGC 2027  
Qy 1801 TTTTCCCTGGGTTCTTATATACAGTTTAAATTTTCTTATGTCCTTCTTATGGCAGTCA 1860  
Db 2028 TTTTCCCTGGGTTCTTATATACAGTTTAAATTTTCTTATGTCCTTCTTATGGCAGTCA 2087  
Qy 1861 TTGCGACAGCTTCTTTGTTATTTCTCCTCAAGTAGCAGCAATTTGTGATATTTCTGCTTTTT 1920  
Db 2088 TTGCGACAGCTTCTTTGTTATTTCTCCTCAAGTAGCAGCAATTTGTGATATTTCTGCTTTTT 2147  
Qy 1921 TCCCTTATGGAATATCATCTGATTTTGTGTTTAAATGCTGACACTCTTTTAAATAAT 1980  
Db 2148 TCCCTTATGGAATATCATCTGATTTTGTGCTTTTAAATGCTGACACTCTTTTAAATAAT 2207  
Qy 1981 CAABAACATGTGGGAATAGTTGAAATTTTGTACTGTGCTTTTGGATTTATTTGGCCTTA 2040  
Db 2208 CAABAACATGTGGGAATAGTTGAAATTTTGTACTGTGCTTTTGGATTTATTTGGCCTTA 2267  
Qy 2041 TGATAATCTCATAGAAAGTTTTCCTCAAAATCGTTAGTGTGGCTTTTTCAGTCTTTTCTGTC 2100  
Db 2268 TGATAATCTCATAGAAAGTTTTCCTCAAAATCGTTAGTGTGGCTTTTTCAGTCTTTTCTGTC 2327  
Qy 2101 ACTGTACTTTTGTGATTTGGTATTTGCACAGCTCATGCAATTTAGAAATTTAATGAGGTG 2160  
Db 2328 ACTGTACTTTTGTGATTTGGTATTTGCACAGCTCATGCAATTTAGAAATTTTAAAGAGGTG 2387  
Qy 2161 CTTTCATTTTCAAAATTTGATGCGAGCCCATATCTCTAATTTAATTAATTAATTAATTAATTAAT 2220  
Db 2388 CTTTCATTTTCAAAATTTGATGCGAGCCCATATCTCTAATTTAATTAATTAATTAATTAATTAAT 2447  
Qy 2221 CACTTAAATAGTATATTTATGTCCTCTTGGCTGTCTATCTTGTATCAAGTCAATTCAGGGG 2280  
Db 2448 CACTTAAATAGTATATTTATGTCCTCTTGGCTGTCTATCTTGTATCAAGTCAATTCAGGGG 2507  
Qy 2281 AATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCTTCATATTTGGTCAAGAGTA 2340  
Db 2508 AATTTGGCTTACGGAGATCATCTTTATATATTTTCTGAAGCTTCATATTTGGTCAAGAGTA 2567

QY 2341 AAAGAAATATGAGGAGTTATCAGAGGCAATGTTAATGAAATATATAGTTTTAGTGAAA 2400  
DB 2568 AAAGAAATATGAGGAGTTATCAGAGGCAATGTTAATGAAATATATAGTTTTAGTGAAA 2627  
QY 2401 TTATGAGCCAGTTCTTCAGAAATTTGAGAAAGAGCCATAGAAATTTAGTGATTC 2460  
DB 2628 TTATGAGCCAGTTCTTCAGAAATTTGAGAAAGAGCCATAGAAATTTAGTGATTC 2687  
QY 2461 AGAAGACATACAGAAAGAGGTCGAAATGTTGAGGCTTTGAGAAATTTGTCATTGACA 2520  
DB 2688 AGAAGACATACAGAAAGAGGTCGAAATGTTGAGGCTTTGAGAAATTTGTCATTGACA 2747  
QY 2521 TATATGAGGTCAGATTAATGTCCTTACTTGGCCACAGTGGAAACAGAGAAAGAGTACATTGA 2580  
DB 2748 TATATGAGGTCAGATTAATGTCCTTACTTGGCCACAGTGGAAACAGAGAAAGAGTACATTGA 2807  
QY 2581 TGAATATCTTTGAGGTCAGATTAATGTCCTTACTTGGCCACAGTGGAAACAGAGAAAGAGTACATTGA 2640  
DB 2808 TGAATATCTTTGAGGTCAGATTAATGTCCTTACTTGGCCACAGTGGAAACAGAGAAAGAGTACATTGA 2867  
QY 2641 GAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATGATTTGGCATTGTCACAGT 2700  
DB 2868 GAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATGATTTGGCATTGTCACAGT 2927  
QY 2701 TAGATATACATTTGATGTTTGAAGCAAGAAATGATTTGGCATTGTCACAGT 2760  
DB 2928 TAGATATACATTTGATGTTTGAAGCAAGAAATGATTTGGCATTGTCACAGT 2987  
QY 2761 AAGGATACAGCAACAAATATAATACAGAAAGTGCAGAGGTTTACTAGATTTAGACA 2820  
DB 2988 AAGGATACAGCAACAAATATAATACAGAAAGTGCAGAGGTTTACTAGATTTAGACA 3047  
QY 2821 TGCAGATCTCAAGATTAACCAAGCTAAATTAAGTGGTGGTCAAAAAGAAAGCTGT 2880  
DB 3048 TGCAGATCTCAAGATTAACCAAGCTAAATTAAGTGGTGGTCAAAAAGAAAGCTGT 3107  
QY 2881 CATTAGGAATGCTGTTCTGGGAACCCAAAGATGCTGCTAGATGAACCAACAGCTG 2940  
DB 3108 CATTAGGAATGCTGTTCTGGGAACCCAAAGATGCTGCTAGATGAACCAACAGCTG 3167  
QY 2941 GAATGGACCCCTGTTCTGACATATGATGGAATCTTTTAAATAACAGAAAGCCAAATC 3000  
DB 3168 GAATGGACCCCTGTTCTGACATATGATGGAATCTTTTAAATAACAGAAAGCCAAATC 3227  
QY 3001 GGGTGACAGTGTTCAGTACTCATTTGATGAGTGAAGTGAATCTTTGCGAGATAGGAAG 3060  
DB 3228 GGGTGACAGTGTTCAGTACTCATTTGATGAGTGAAGTGAATCTTTGCGAGATAGGAAG 3287  
QY 3061 CTGTGATATCAAGAGGAATGCTGAAATGTTGTTGTTCTTCAATGTTCTTCAAAAGTAAAT 3120  
DB 3288 CTGTGATATCAAGAGGAATGCTGAAATGTTGTTGTTCTTCAATGTTCTTCAAAAGTAAAT 3347  
QY 3121 GGGGGATCGGCTACCGCTGAGCATGATACAGAAATATGTTGCCACAGAAATCTCTTT 3180  
DB 3348 GGGGGATCGGCTACCGCTGAGCATGATACAGAAATATGTTGCCACAGAAATCTCTTT 3407  
QY 3181 CTTTCAGTGGTTAAACACATATACCTGGAGTACTTTTATACACAGAAATGACCAACAC 3240  
DB 3408 CTTTCAGTGGTTAAACACATATACCTGGAGTACTTTTATACACAGAAATGACCAACAC 3467  
QY 3241 TTGTGATAGCTTCCCTTTCAAGACATGGAACAAATTTTCAAGTTGTTTCTGCCCTAG 3300  
DB 3468 TTGTGATAGCTTCCCTTTCAAGACATGGAACAAATTTTCAAGTTGTTTCTGCCCTAG 3527  
QY 3301 ACAGTCAATCAAAATTTGGGTGTCATTTCTTATGTTGTTTCCATGACGACTTTGGAAGACG 3360  
DB 3528 ACAGTCAATCAAAATTTGGGTGTCATTTCTTATGTTGTTTCCATGACGACTTTGGAAGACG 3587  
QY 3361 TATTTTTAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATTTATAGTATTTACTC 3420  
DB 3588 TATTTTTAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATTTATAGTATTTACTC 3647

QY 3421 AGCAGCCACTGGAGGAAGAAATGGAATTTCAAAATCTTTTGTGATGAAATGGAACAGAGCTTAC 3480  
DB 3648 AGCAGCCACTGGAGGAAGAAATGGAATTTCAAAATCTTTTGTGATGAAATGGAACAGAGCTTAC 3707  
QY 3481 TTATTTCTTTCTGAAACCAAGGCTTCTAGTGAAGCACCATTGAGCCCTTTGGAAACACAGA 3540  
DB 3708 TTATTTCTTTCTGAAACCAAGGCTTCTAGTGAAGCACCATTGAGCCCTTTGGAAACACAGA 3767  
QY 3541 TGTATACAAATAGCAAAAGTTTCTTTTCTTTTACCTTTGAAACGTAAGAAAGTAAATCAGTGAGAT 3600  
DB 3768 TGTATACAAATAGCAAAAGTTTCTTTTCTTTTACCTTTGAAACGTAAGAAAGTAAATCAGTGAGAT 3827  
QY 3601 CAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTTGCTTCATC 3660  
DB 3828 CAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTTGCTTCATC 3887  
QY 3661 ACTCTTTTAAATAGTGTGTTCCCATCAAACTGTTCCAGACTTATATTTCTTAAAC 3720  
DB 3888 ACTCTTTTAAATAGTGTGTTCCCATCAAACTGTTCCAGACTTATATTTCTTAAAC 3947  
QY 3721 CTGGAGACAAACCAACATAAATACAAAACAAAGTCTGCTTCTTCAAAATTTCTGCTGACTCAG 3780  
DB 3948 CTGGAGACAAACCAACATAAATACAAAACAAAGTCTGCTTCTTCAAAATTTCTGCTGACTCAG 4007  
QY 3781 ATATCAGTGATCTTATTTAGCTTTTTCACAAAGCCAGAACATAATGTTGTCAGATGATTAATG 3840  
DB 4008 ATATCAGTGATCTTATTTAGCTTTTTCACAAAGCCAGAACATAATGTTGTCAGATGATTAATG 4067  
QY 3841 ACAGTGACTATGATCCGTTGCTCCCATAGTGGGCTTTAAATGTTGATGATGATTCAGAAA 3900  
DB 4068 ACAGTGACTATGATCCGTTGCTCCCATAGTGGGCTTTAAATGTTGATGATGATTCAGAAA 4127  
QY 3901 AGGACTATGTTTTGAGCTGTTTTCAACAGTACTATGTTTATCTTTACCTATATTAG 3960  
DB 4128 AGGACTATGTTTTGAGCTGTTTTCAACAGTACTATGTTTATCTTTACCTATATTAG 4187  
QY 3961 TGAATATCATTTAGTAACTACTATCTTTTATCAATTTAAATGTCAGTGAACCATCCAGATCT 4020  
DB 4188 TGAATATCATTTAGTAACTACTATCTTTTATCAATTTAAATGTCAGTGAACCATCCAGATCT 4247  
QY 4021 GGAGTACCCCATCTTTTCAAGAAATTAATGATAGTTTTTAAATGAGCTGATTTTTC 4080  
DB 4248 GGAGTACCCCATCTTTTCAAGAAATTAATGATAGTTTTTAAATGAGCTGATTTTTC 4307  
QY 4081 AAGCAGCTTGTCTGGAATCATTTGTTACTGCAATGCCACTTACTTGGCCATGAAAATG 4140  
DB 4308 AAGCAGCTTGTCTGGAATCATTTGTTACTGCAATGCCACTTACTTGGCCATGAAAATG 4367  
QY 4141 CAGAGAAATCATAGATCAAGCTTATTAATCAACTTAAACCTTTTCCAGGCTTTTGGCCATCTG 4200  
DB 4368 CAGAGAAATCATAGATCAAGCTTATTAATCAACTTAAACCTTTTCCAGGCTTTTGGCCATCTG 4427  
QY 4201 CATATTTGGAATGGAACAGCTGTTGTTGATATCCCTTTATTTTTTATCATCTTTATTTGA 4260  
DB 4428 CATATTTGGAATGGAACAGCTGTTGTTGATATCCCTTTATTTTTTATCATCTTTATTTGA 4487  
QY 4261 TGCCTAGAGAGCTTACTGGAATCATTTTATGGAATATATTTTATATCTGTAAGTCCCTG 4320  
DB 4488 TGCCTAGAGAGCTTACTGGAATCATTTTATGGAATATATTTTATATCTGTAAGTCCCTG 4547  
QY 4321 CTGTGTTTTTGGCTTATTTGGTTATGTTCCCATCAGTTATTTGTTTCACTTATATTTGCTT 4380  
DB 4548 CTGTGTTTTTGGCTTATTTGGTTATGTTCCCATCAGTTATTTGTTTCACTTATATTTGCTT 4607  
QY 4381 CTTTCACCTTTTAAAGAAAATTTTAAATAACAAAGAAATTTTGGTCAATTTATCTATCTGTGG 4440  
DB 4608 CTTTCACCTTTTAAAGAAAATTTTAAATAACAAAGAAATTTTGGTCAATTTATCTATCTGTGG 4667  
QY 4441 CAGCGTTGCTGTTGATGCAATCAGTGAATTAATCTTTTATGGAATACAGATTTGCAA 4500  
DB 4668 CAGCGTTGCTGTTGATGCAATCAGTGAATTAATCTTTTATGGAATACAGATTTGCAA 4727  
QY 4501 CTATTTCTTCAATGCTTTTGTATCATCATTTCCAATCTATCCAATCTTAGGTTGCTGTA 4560

Db	4728	 CTATCTTCATATATGCGCCTTTTGTATCATCATCAATCTATCCACTTCTAGGTGCTGA	4787
Qy	4561	TTTCTTTTCATAAGATTTCTTTGGAGAAATGTCAGAAAAATGTGGACACCTATAATCCAT	4620
Db	4788	TTTCTTTTCATAAGATTTCTTTGGAGAAATGTCAGAAAAATGTGGACACCTATAATCCAT	4847
Qy	4621	GGGATAGGCTTTTCAGTAGCTGTTATATCGCCTTTACCTGCGAGTGTGTACTGTGGAATTTCC	4680
Db	4848	GGGATAGGCTTTTCAGTAGCTGTTATATCGCCTTTACCTGCGAGTGTGTACTGTGGAATTTCC	4907
Qy	4681	TCCTTCAATATCTATGAGAAAAATATGGAGGCGAGATCAATTAAGAAAAAGATCCCTTTTCA	4740
Db	4908	TCCTTCAATATCTATGAGAAAAATATGGAGGCGAGATCAATTAAGAAAAAGATCCCTTTTCA	4967
Qy	4741	GAACCTTTTCAACGAACTTAAATAAGGAAGCTTCCAGAACCAACCAGACAATGAGGATG	4800
Db	4968	GAACCTTTTCAACGAACTTAAATAAGGAAGCTTCCAGAACCAACCAGACAATGAGGATG	5027
Qy	4801	AAGATGAAGATGTCAAAAGCTGAAAGACTAAAGGTCAAAAGAGCTGATGGGTTCGCAGTGT	4860
Db	5028	AAGATGAAGATGTCAAAAGCTGAAAGACTAAAGGTCAAAAGAGCTGATGGGTTCGCAGTGT	5087
Qy	4861	GTGAGGAGAAACCATTCATTTATGGTCAGCAATTTGCAATAAGAAATATGATGACAAGAAAG	4920
Db	5088	GTGAGGAGAAACCATTCATTTATGGTCAGCAATTTGCAATAAGAAATATGATGACAAGAAAG	5147
Qy	4921	ATTTTCTTTTCAAGAAAAAGTAAAGAGTGGCACTAAATACATCTCTTTCTGTGTGA	4980
Db	5148	ATTTTCTTTTCAAGAAAAAGTAAAGAGTGGCACTAAATACATCTCTTTCTGTGTGA	5207
Qy	4981	AAAAAGGAGAGATCTTAGGACTATTGGGTCCAAATGTGCTGGCAAAAGCACAAATATTA	5040
Db	5208	AAAAAGGAGAGATCTTAGGACTATTGGGTCCAAATGTGCTGGCAAAAGCACAAATATTA	5267
Qy	5041	ATATTTCTGGTTGGTGATTTGAACCAACTTCAGGCCAGGTATTTTATAGGAGATTAATCTTT	5100
Db	5268	ATATTTCTGGTTGGTGATTTGAACCAACTTCAGGCCAGGTATTTTATAGGAGATTAATCTTT	5327
Qy	5101	CAGAGACAAGTGAAGATGATGATTCACCTGAAGTGTATGGGTACTGTCCTCAGATAAAC	5160
Db	5328	CAGAGACAAGTGAAGATGATGATTCACCTGAAGTGTATGGGTACTGTCCTCAGATAAAC	5387
Qy	5161	CTTTGTGGCCAGATPACTACATTCGAGAAACATTTTGAATTTATGGAGCTGTCAAGAGAA	5220
Db	5388	CTTTGTGGCCAGATPACTACATTCGAGAAACATTTTGAATTTATGGAGCTGTCAAGAGAA	5447
Qy	5221	TGAGTGCAAGTGACATGAAGAAAGTCTAAGTCGAATAACACATGCACTTGATTTAAAG	5280
Db	5448	TGAGTGCAAGTGACATGAAGAAAGTCTAAGTCGAATAACACATGCACTTGATTTAAAG	5507
Qy	5281	AACATCTTCAGAGACTGTAAAGAACTACTGCGAGGAATCAACAGAAAGTTGTGTTTTG	5340
Db	5508	AACATCTTCAGAGACTGTAAAGAACTACTGCGAGGAATCAACAGAAAGTTGTGTTTTG	5567
Qy	5341	CTCTAAGTATGCTAGGGAATCCTCAGATTAATTTTGTCTAGATGACCACTACAGGTATGG	5400
Db	5568	CTCTAAGTATGCTAGGGAATCCTCAGATTAATTTTGTCTAGATGACCACTACAGGTATGG	5627
Qy	5401	ATCCAAAGGCCAAACAGCACATGTGGCGAGCAATTCGAACTGCAATTTAAAAACAGAAAGC	5460
Db	5628	ATCCAAAGGCCAAACAGCACATGTGGCGAGCAATTCGAACTGCAATTTAAAAACAGAAAGC	5687
Qy	5461	GGGCTGCTATTCGACCACTCATATATGAGGAGGCGAGAGGCTGTCTGTGATCGAGTAG	5520
Db	5688	GGGCTGCTATTCGACCACTCATATATGAGGAGGCGAGAGGCTGTCTGTGATCGAGTAG	5747
Qy	5521	CTATCATGGTGTCTGGGCGAGTTAAGATGTATCGGAAACAGTACAACTCTAAAGAGTAAAT	5580
Db	5748	CTATCATGGTGTCTGGGCGAGTTAAGATGTATCGGAAACAGTACAACTCTAAAGAGTAAAT	5807
Qy	5581	TTGGAAAAGGCTACTTTTTTGGAAATTAATTTGAAGGACTCGATAGAAAACTAGAAAGTAG	5640

[illegible]

PI	Chen H, Kilinski L, Le Bihan S;
XX	WPI: 2003-066798/06.
DR	P-PSDB; ABG72424.
XX	
PT	Novel isolated ATP binding cassette transporter family polypeptide,
PT	ABCAs, useful for treating disorders associated with aberrant or
PT	unwanted ABCAs transporter expression or activity -
XX	Claim 2; Page 39-42; 52pp; English.
PS	
XX	
CC	The invention describes an isolated ATP binding cassette (ABC)
CC	transporter family polypeptide (I), designated ABCA5. (I) or the
CC	polynucleotide encoding it (II) are useful as targets for developing
CC	modulating agents to regulate a variety of cellular processes,
CC	particularly the transport of neurotoxic molecules, e.g., beta-amyloid
CC	peptide (Abeta), across cell membranes or, e.g., the blood brain barrier
CC	(BBB), as targets for developing modulating agents of multi-drug
CC	resistance, as diagnostic and therapeutic tools, or to treat disorders
CC	associated with aberrant or unwanted ABCAs transporter expression or
CC	activity. (I), (II) or a host cell (III) expressing (II) are useful in
CC	screening assays, detection assays (e.g., chromosomal mapping, tissue
CC	typing, forensic biology), predictive medicine (e.g., diagnostic assays,
CC	prognostic assays, monitoring clinical trials and pharmacogenomics), and
CC	in methods of treatment (e.g., therapeutic and prophylactic). (I) or
CC	(III) are useful as reagents or targets in assays applicable to treatment
CC	and diagnosis of ABCA5-mediated or related disorders. (I) Is useful to
CC	screen for naturally occurring ABCA5 substrates; to screen for drugs
CC	or compounds which modulate ABCA5 activity; as a bait protein in a yeast
CC	two-hybrid or three-hybrid assay; and to identify other proteins which
CC	bind to or interact with ABCA5. (II) Is useful in: gene therapy; to
CC	detect ABCA5 mRNA or a genetic alteration in a ABCA5 gene; to modulate
CC	ABCA5 activity; to locate gene regions associated with genetic disease
CC	or to associate ABCA5 with the disease, to identify an individual from
CC	a minute biological sample (tissue typing), and to aid in forensic
CC	identification of the biological sample. This sequence encodes a
CC	novel human ATP binding cassette (ABC) A5 transporter.
XX	
SQ	Sequence 5475 BP; 1705 A; 917 C; 1048 G; 1804 T; 1 other;
	Query Match 82.3%; Score 5373-2; DB 25; Length 5475;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 5375; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY	994 AGGTATTTTCAGAAACATGCTCCACTGCATAATTAGGAGGTAGGAGTTGGAGACAGACCA 1053
DB	
DB	97 AGCTTATTCAGAAACATGCTCCACTGCATAATTAGGAGGTAGGAGTTGGAGACAGACCA 156
QY	1054 GAACACTTCTACTGAAGAATTACTTAATTAATGACAGAACCAAAAGAGTAGTGTCAGG 1113
DB	
DB	157 GAACACTTCTACTGAAGAATTACTTAATTAATGACAGAACCAAAAGAGTAGTGTCAGG 216
QY	1114 AAATCTCTTTTCCACTATTTTTTTTATTTTGGTTAATTAATTAATGACATGATGCCAA 1173
DB	
DB	217 AAATCTCTTTTCCACTATTTTTTTTATTTTGGTTAATTAATTAATGACATGATGCCAA 276
QY	1174 ATAGAATAATGAAGAAGTGCCTAATATAGAATCAATCCTATGGACAAGTTTACTCTTT 1233
DB	
DB	277 ATAAGAATAATGAAGAAGTGCCTAATATAGAATCAATCCTATGGACAAGTTTACTCTTT 336
QY	1234 CTAAATCTAATCTTGGATATATCTCAGTGACTAATATTAACAGCAGCATCATGCGAAAAG 1293
DB	
DB	337 CTAAATCTAATCTTGGATATATCTCAGTGACTAATATTAACAGCAGCATCATGCGAAAAG 396
QY	1294 TGCTCTACTCATCTACTCCTGATGTCATTAATTAATGACATATACAAATGAAAAAGAAA 1353
DB	
DB	397 TGCTCTACTCATCTACTCCTGATGTCATTAATTAATGACATATACAAATGAAAAAGAAA 456
QY	1354 TGTTAACATCCAGTCTCTCTAAGCCGAGCACTTTGTAGGTGGTTTCAAAGACTCCA 1413
DB	
DB	457 TGTTAACATCCAGTCTCTCTAAGCCGAGCACTTTGTAGGTGGTTTCAAAGACTCCA 516
QY	1414 TGTCCTPATGAACCTCGTTTTTTTCTCGATATGATTCAGATATCTTCTATTTATATGGATT 1473

1597 AGGCTTTGAGAAATTTGTCATTTGACATATATGAGGCTCAGATTACTGCCCTTACTTTGGCC 1656  
2554 ACAGTGAACAGGAAAGTACATTTGATGATATCTTTTGTGACCTCTGCCACCTTCTG 2613  
1657 ACAGTGAACAGGAAAGTACATTTGATGATATCTTTTGTGACCTCTGCCACCTTCTG 1716  
2614 ATGGGTTTGATCTATATATGGAACACAGAGCTCTCAGAAATAGATGAAATGTTTGAAGCAA 2673  
1717 ATGGGTTTGATCTATATATGGAACACAGAGCTCTCAGAAATAGATGAAATGTTTGAAGCAA 1776  
2674 GAAAAATGATTTGGCAATTTGCTCCAGTTAGATATACACTTTGATGTTTGAAGTGAAG 2733  
1777 GAAAAATGATTTGGCAATTTGCTCCAGTTAGATATACACTTTGATGTTTGAAGTGAAG 1836  
2734 AAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCAACATATATATACAGAG 2793  
1837 AAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCAACATATATATACAGAG 1896  
2794 TGCAAGAGGTTTACTAGATTTAGACATGACAGTATCAAGAGTAAACCAAGCTTAAAAAT 2853  
1897 TGCAAGAGGTTTACTAGATTTAGACATGACAGTATCAAGAGTAAACCAAGCTTAAAAAT 1956  
2854 TAAGTGTGTGTCAAAAAGAAAGCTGTCATTTAGGAATGCTGTTCTTTGGGAAACCCAAAGA 2913  
1957 TAAGTGTGTGTCAAAAAGAAAGCTGTCATTTAGGAATGCTGTTCTTTGGGAAACCCAAAGA 2016  
2914 TACTGCTGTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATTTGTATGGA 2973  
2017 TACTGCTGTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATTTGTATGGA 2076  
2974 ATCTTTTAAAAATCAGAAAAAGCCAAATCGGTGACAGTGTTCAGTACTCATTTTCATGGATG 3033  
2077 ATCTTTTAAAAATCAGAAAAAGCCAAATCGGTGACAGTGTTCAGTACTCATTTTCATGGATG 2136  
3034 AAGCTGACATTTCTTGCAGATAGGAAGCTGTGTATATCAAGGAATGCTGAAATGTGTTG 3093  
2137 AAGCTGACATTTCTTGCAGATAGGAAGCTGTGTATATCAAGGAATGCTGAAATGTGTTG 2196  
3094 GTTCTTCAATGTTCTTCAAAAGTAAATGGGGATGCGCTACCGCTGAGCATGTACATAG 3153  
2197 GTTCTTCAATGTTCTTCAAAAGTAAATGGGGATGCGCTACCGCTGAGCATGTACATAG 2256  
3154 ACAAAATATGTGACACAGAAATCTTTCTTCACTGGTTTAAACACATATACCTGGAGCTA 3213  
2257 ACAAAATATGTGACACAGAAATCTTTCTTCACTGGTTTAAACACATATACCTGGAGCTA 2316  
3214 CTTTATTAACACAGAAATGACCAACTGTGTATAGCTTTGCCCTTCAAGGACATGGACA 3273  
2317 CTTTATTAACACAGAAATGACCAACTGTGTATAGCTTTGCCCTTCAAGGACATGGACA 2376  
3274 AATTTTCAGGTTTGTGTTCTGCCCTAGACAGTCAATTCAAATTTGGGTGTCATTTCTTATG 3333  
2377 AATTTTCAGGTTTGTGTTCTGCCCTAGACAGTCAATTCAAATTTGGGTGTCATTTCTTATG 2436  
3334 GTGTTTCCATGACGACTTTTGAAGACGCTATTTTAAAGCTAGAAAGTTGAAGCAGAAATG 3393  
2437 GTGTTTCCATGACGACTTTTGAAGACGCTATTTTAAAGCTAGAAAGTTGAAGCAGAAATG 2496  
3394 ACCAAGCAGATATATAGTATTTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAAT 3453  
2497 ACCAAGCAGATATATAGTATTTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAAT 2556  
3454 CTTTTTGATGAATGGAACAGAGCTTACTTATCTTTCTGAAACCAAGGCTTCTTAGTGA 3513  
2557 CTTTTTGATGAATGGAACAGAGCTTACTTATCTTTCTGAAACCAAGGCTTCTCTAGTGA 2616  
3514 GCACCATGAGCCTTTTGAAGACGCTATATATCTTCTGAAACCAAGGCTTCTCTACCT 3573  
2617 GCACCATGAGCCTTTTGAAGACGCTATATATCTTCTGAAACCAAGGCTTCTCTACCT 2676  
3574 TGAACGCTGAAGTAAATCAGTGAATCAGTGTGCTTCTGCTTTAAATTTTTCACAG 3633  
2677 TGAACGCTGAAGTAAATCAGTGAATCAGTGTGCTTCTGCTTTAAATTTTTCACAG 2736

3634 TTCCAGATTTTATGTTTGTGTTTGTGTTTCACTCTTTTAAAAATGCTGTGTTTCCCAATCAAC 3693  
2737 TTCCAGATTTTATGTTTGTGTTTGTGTTTCACTCTTTTAAAAATGCTGTGTTTCCCAATCAAC 2796  
3694 TTGTTTCCAGACTTATATTTTCTAAAACTCGAGAGCAAAACCAATATAAATACAAAAAAGTCT 3753  
2797 TTGTTTCCAGACTTATATTTTCTAAAACTCGAGAGCAAAACCAATATAAATACAAAAAAGTCT 2856  
3754 TGCCTTCTTCAAAATTTCTGCTGACTCAGATATACAGTATCTTATAGCTTTTTCACAAAGC 3813  
2857 TGCCTTCTTCAAAATTTCTGCTGACTCAGATATACAGTATCTTATAGCTTTTTCACAAAGC 2916  
3814 AGAACATATATGTTGACGATGATTAATGACAGTACTATGATATCCCTGGCTCCCATAGTG 3873  
2917 AGAACATATATGTTGACGATGATTAATGACAGTACTATGATATCCCTGGCTCCCATAGTG 2976  
3874 CGGCTTTAAATGTGATGCAATTCAGAAAAAGGACTATGTTTTTGGCAGCTGTTTTCAACAGTA 3933  
2977 CGGCTTTAAATGTGATGCAATTCAGAAAAAGGACTATGTTTTTGGCAGCTGTTTTCAACAGTA 3036  
3934 CTATGGTTTATTTTACTATATATAGTGAATATCAATAGTAACTACTATCTTTTATCAAT 3993  
3037 CTATGGTTTATTTTACTATATATAGTGAATATCAATAGTAACTACTATCTTTTATCAAT 3096  
3994 TAAATGTGACTGAAACCAATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATTAATCTGATA 4053  
3097 TAAATGTGACTGAAACCAATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATTAATCTGATA 3156  
4054 TAGTTTTTAAAAATGTGATGTTTTTCAAGCAGCTTTGCTTGGAAATCAATGTTTACTGCA 4113  
3157 TAGTTTTTAAAAATGTGATGTTTTTCAAGCAGCTTTGCTTGGAAATCAATGTTTACTGCA 3216  
4114 TGCACCTTACTTTTGGCATGGAATGACAGAAATCAAGATCAAGCTTATATCTCAAC 4173  
3217 TGCACCTTACTTTTGGCATGGAATGACAGAAATCAAGATCAAGCTTATATCTCAAC 3276  
4174 TTTAAACTTTTCAAGTCTTTTCCATCTGATATTTGATTTGGCAAGCTGTTGTTGATATCC 4233  
3277 TTTAACTTTTCAAGTCTTTTCCATCTGATATTTGATTTGGCAAGCTGTTGTTGATATCC 3336  
4234 CTTATTTTATATCAATCTTATTTTGTAGTCTAGGAAGCTTACTGGCAATTCATTTATGAT 4293  
3337 CTTATTTTATATCAATCTTATTTTGTAGTCTAGGAAGCTTATTTGCAATTTCAATTTATGAT 3396  
4294 TATATTTTATATCTGTAAGTTCCTTCTGCTGTTTGTGCTTATTTGCTTATGTTTATGTTCCAT 4353  
3397 TATATTTTATATCTGTAAGTTCCTTCTGCTGTTTGTGCTTATTTGCTTATGTTTATGTTCCAT 3456  
4354 CAGTTTATTTCTGTTTCACTTATATTTGCTTCTTCCACCTTTAAGAAAAATTTTAAAAATACCAAG 4413  
3457 CAGTTTATTTCTGTTTCACTTATATTTGCTTCTTCCACCTTTAAGAAAAATTTTAAAAATACCAAG 3516  
4414 AATTTTGGTCAATTTATCTATCTGTCGAGCGTGTGCTTGTATTTGCAATCACTGAAATAA 4473  
3517 AATTTTGGTCAATTTATCTATTTCTGTCGAGCGTGTGCTTGTATTTGCAATCACTGAAATAA 3576  
4474 CTTTCTTTATGGATACACAAATTTGCAATTTCTTCAATTTATGCTTATGCTTATCATCATTC 4533  
3577 CTTTCTTTATGGATACACAAATTTGCAATTTCTTCAATTTATGCTTATGCTTATCATCATTC 3636  
4534 CAATCTATCTCTCTTCTAGGTTGCTTATTTCTTCTTCAATAAAGATTTCTTGGAAAGATGTAC 4593  
3637 CAATCTATCTCTCTTCTAGGTTGCTTATTTCTTCTTCAATAAAGATTTCTTGGAAAGATGTAC 3696  
4594 GAAAAATGTGACACCTTATATATCCATGGATAGGCTTTTCACTAGCTGTTTATGCTTATGCTCTT 4653  
3697 GAAAAATGTGACACCTTATATATCCATGGATAGGCTTTTCACTAGCTGTTTATGCTTATGCTCTT 3756  
4654 ACCTGAGTGTGCTGCTGCTGCTTCTTCTTCAATACTATGAGAAAAATATGAGGCA 4713  
3757 ACCTGAGTGTGCTGCTGCTGCTTCTTCTTCAATACTATGAGAAAAATATGAGGCA 3816

QY 4714 GATCAATAAGAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGCTCTAAAAATAGGAAGC 4773  
DB 3817 GATCAATAAGAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGCTCTAAAAATAGGAAGC 3876  
QY 4774 TTCAGAACCAACAGCAATAGAGATGAAGATGAAGATGTCAGAGCTGGAAGACTAAGG 4833  
DB 3877 TTCAGAACCAACAGCAATAGAGATGAAGATGAAGATGTCAGAGCTGGAAGACTAAGG 3936  
QY 4834 TCAAGAGCTGATGGGTGCGCAGTGTGTGAGAGAAACCATCCATTTATGGTCAGCAAT 4893  
DB 3937 TCAAGAGCTGATGGGTGCGCAGTGTGTGAGAGAAACCATCCATTTATGGTCAGCAAT 3996  
QY 4894 TGCATAAAGATATGATGACAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGG 4953  
DB 3997 TGCATAAAGATATGATGACAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGG 4056  
QY 4954 CAACTAAATACATCTCTTCTGTCGTAAGAAAGAGAGATCTTAGGACTATTTGGGTCCAA 5013  
DB 4057 CAACTAAATACATCTCTTCTGTCGTAAGAAAGAGAGATCTTAGGACTATTTGGGTCCAA 4116  
QY 5014 ATGGTGTGTCGCAAAAGCACAAATTAATAATTTCTGGTGTGGTGATATTGAACCAACTTCAG 5073  
DB 4117 ATGGTGTGTCGCAAAAGCACAAATTAATAATTTCTGGTGTGGTGATATTGAACCAACTTCAG 4176  
QY 5074 GCCAGGTATTTTGGAGATATTCTTTCAGAGACAAGTGAAGATGATTCACCTGAAGT 5133  
DB 4177 GCCAGGTATTTTGGAGATATTCTTTCAGAGACAAGTGAAGATGATTCACCTGAAGT 4236  
QY 5134 GTATGGGTACTCTCTCAGATAAACCCCTTTGTGGCCAGATACATATTCAGAGAACAT 5193  
DB 4237 GTATGGGTACTCTCTCAGATAAACCCCTTTGTGGCCAGATACATATTCAGAGAACAT 4296  
QY 5194 TTGAAATTTATGGAGCTGTCAAAAGGAATGAGTCAAGTGCATGAAAGAGTCAATAGTC 5253  
DB 4297 TTGAAATTTATGGAGCTGTCAAAAGGAATGAGTGCATGAAAGAGTCAATAGTC 4356  
QY 5254 GAATAACATGACATCTGATTTTAAAGAACATCTTCAGAGACTGTAAAGAAACTACCTG 5313  
DB 4357 GAATAACATGACATCTGATTTTAAAGAACATCTTCAGAGACTGTAAAGAAACTACCTG 4416  
QY 5314 CAGGAATCAACCAAGATGTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTAATT 5373  
DB 4417 CAGGAATCAACCAAGATGTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTAATT 4476  
QY 5374 TGTAGATGAACCATCTACAGGTATGATCCCAAGCCAAACAGACATGTGGCGAGCAA 5433  
DB 4477 TGTAGATGAACCATCTACAGGTATGATCCCAAGCCAAACAGACATGTGGCGAGCAA 4536  
QY 5434 TTCGAATCGCATTTAAAAACAGAAAGCGGCTCTATTCTGACCCTCACTATATGGAGG 5493  
DB 4537 TTCGAATCGCATTTAAAAACAGAAAGCGGCTCTATTCTGACCCTCACTATATGGAGG 4596  
QY 5494 AGCAGAGGCTGTCTGTGATCGAGTAGCTATCATGTGTCTGGGAGTTAAGATGTATCG 5553  
DB 4597 AGCAGAGGCTGTCTGTGATCGAGTAGCTATCATGTGTCTGGGAGTTAAGATGTATCG 4656  
QY 5554 GAACAGTACAACTCTAAAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTTGA 5613  
DB 4657 GAACAGTACAACTCTAAAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTTGA 4716  
QY 5614 AGGACTGGATAGAAAACCTAGAAAGTAGACCGCCCTTCAAGAGAGAAATTCAGTATATTTCC 5673  
DB 4717 AGGACTGGATAGAAAACCTAGAAAGTAGACCGCCCTTCAAGAGAGAAATTCAGTATATTTCC 4776  
QY 5674 CAAATGCAAGCCGTGAGAAAGTTTCTCTATTTTGGCTTATAAAATTCCTAAGGAAG 5733  
DB 4777 CAAATGCAAGCCGTGAGAAAGTTTCTCTATTTTGGCTTATAAAATTCCTAAGGAAG 4836  
QY 5734 ATGTTTCAGTCCCTTTCACAACTCTTTTAAAGCTGGAAGCTAAACATGCTTTTGCCA 5793  
DB 4837 ATGTTTCAGTCCCTTTCACAACTCTTTTAAAGCTGGAAGCTAAACATGCTTTTGCCA 4896  
QY 5794 TTGAAGAATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTTGTAGAACTCACTAAAG 5853

DB 4897 TTGAAGAATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTTGTAGAACTCACTAAAG 4956  
QY 5854 AACAGAGAGAGGAAGATAATAGTTGTGGAACTTTTAAACAGCACACTTTTGTGGGAACGAA 5913  
DB 4957 AACAGAGAGAGGAAGATAATAGTTGTGGAACTTTTAAACAGCACACTTTTGTGGGAACGAA 5016  
QY 5914 CACAAGAGATAGAGTAGTATTTTGAATTTGTATTTCTGGTCTGCTTACTTGGGACTTCT 5973  
DB 5017 CACAAGAGATAGAGTAGTATTTTGAATTTGTATTTCTGGTCTGCTTACTTGGGACTTCT 5076  
QY 5974 TTCTTTTTCACCTAATTTTAACTTTTGGTTTAAAAAGTTTTTATTGGAATGGTAACTGGA 6033  
DB 5077 TTCTTTTTCACCTAATTTTAACTTTTGGTTTAAAAAGTTTTTATTGGAATGGTAACTGGA 5136  
QY 6034 GAACCAAGAACCCACCTTGAAATTTTCTAAGCTCTTAAATTTGAAATGCTGTGGTGTGTG 6093  
DB 5137 GAACCAAGAACCCACCTTGAAATTTTCTAAGCTCTTAAATTTGAAATGCTGTGGTGTGTG 5196  
QY 6094 TTTTGTCTTTCTTAAATAAAACGTATGTATTAATTAAGTGAAGCTGCATGTTTGTATTGA 6153  
DB 5197 TTTTGTCTTTCTTAAATAAAACGTATGTATTAATTAAGTGAAGCTGCATGTTTGTATTGA 5256  
QY 6154 AGTATATTGAACTATATAGTTTGTATGTATCATCTTTTTCACCATTCAGAAACAGTGTCT 6213  
DB 5257 AGTATATTGAACTATATAGTTTGTATGTATCATCTTTTTCACCATTCAGAAACAGTGTCT 5316  
QY 6214 GAATTTGTGATTTTAAAGGAATTTGTAATAGATAGTTTATTATTAAAGTTTCTTAAAGTT 6273  
DB 5317 GAATTTGTGATTTTAAAGGAATTTGTAATAGATAGTTTATTATTAAAGTTTCTTAAAGTT 5376  
QY 6274 TATGCCATCTTCTTAAATAAGTACGTAATGTTCCCAATCTAAATAAAAACTAATACATAA 6333  
DB 5377 TATGCCATCTTCTTAAATAAGTACGTAATGTTCCCAATCTAATATAAAAACTAATACATAA 5436  
QY 6334 CTAAATGATAGAAAAAGATACATAAAGCAATGTGAAAGTT 6372  
DB 5437 CTAAATGATAGAAAAAGATACATAAAGCAATGTGAAAGTT 5475

RESULT 4  
ABS57749  
ID ABS57749 standard; cDNA; 5463 BP.  
XX  
AC ABS57749;  
XX  
DT 04-FEB-2003 (first entry)  
XX  
DE cDNA encoding novel human ATP binding cassette ABCA5 transporter #1.  
XX  
KW Human; ATP binding cassette; ABC; ABCA5; transporter;  
KW neurotoxin transport; beta-amyloid peptide; chromosome mapping;  
KW blood brain barrier transport; tissue typing; predictive medicine;  
KW ABCA5 mediated disorder; ABCA5 related disorder; gene therapy;  
KW gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT CDS Location/Qualifiers  
FT 114..5030  
FT /\*tag= a  
FT /product= "ABCA5"  
FT /note= "ATP binding cassette (ABC) A5 transporter"  
FT /transl\_except= (pos:3273..3275, aa:Ser)  
XX  
US2002123107-A1.  
XX  
PD 05-SEP-2002.  
XX  
PF 01-MAR-2002; 2002US-0090458.  
XX  
PR 02-MAR-2001; 2001US-272885P.  
XX







QY	2494	AGGCTTTGAGAAATTTGTCATTTGACATATATAGGGGTGAGATTAAGCTTACTTGCC	2553
Db	1597	AGGCTTTGAGAAATTTGTCATTTGACATATATAGGGGTGAGATTAAGCTTACTTGCC	1656
QY	2554	ACAGTGAACAGGAAGTACATTGATGAATATCTTTTGGAGCTCTGCCACCTTCTG	2613
Db	1657	ACAGTGAACAGGAAGTACATTGATGAATATCTTTTGGAGCTCTGCCACCTTCTG	1716
QY	2614	ATGGGTTTGCAATCTATATATGGAACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAA	2673
Db	1717	ATGGGTTTGCAATCTATATATGGAACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAA	1776
QY	2674	GAAAAATGATTTGGCAATTTGCCACAGTTAGATATACACTTTTGGATTTTGCAGAGTGAAG	2733
Db	1777	GAAAAATGATTTGGCAATTTGCCACAGTTAGATATACACTTTTGGATTTTGCAGAGTGAAG	1836
QY	2734	AAAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAACATATATACNAGAAG	2793
Db	1837	AAAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAACATATATATACNAGAAG	1896
QY	2794	TGCAGAAAGGTTTACTAGATTTAGACATGCAGACTATCAAGATTAACCAAGCTAAAAAAT	2853
Db	1897	TGCAGAAAGGTTTACTAGATTTAGACATGCAGACTATCAAGATTAACCAAGCTAAAAAAT	1956
QY	2854	TAAGTGTGTGCAAAAAAGAAAGCTGTCAATTAGGAATTTGCTTTGGGAACCCAAAGA	2913
Db	1957	TAAGTGTGTGCAAAAAAGAAAGCTGTCAATTAGGAATTTGCTTTGGGAACCCAAAGA	2016
QY	2914	TACTGTGCTAGATGAACCAACAGCTGGAATGGAACCTGTTCTCGACATATTTGTATGGA	2973
Db	2017	TACTGTGCTAGATGAACCAACAGCTGGAATGGAACCTGTTCTCGACATATTTGTATGGA	2076
QY	2974	ATCTTTTAAAAATACAGAAAAAGCCAACTCGGTGACAGTGTTCAGTACTCATTTTCATGATG	3033
Db	2077	ATCTTTTAAAAATACAGAAAAAGCCAACTCGGTGACAGTGTTCAGTACTCATTTTCATGATG	2136
QY	3034	AAGCTGACATTTCTTGACAGATAGGAAGCTGTGTATATACAAAGGAATGCTGAAATGTGTTG	3093
Db	2137	AAGCTGACATTTCTTGACAGATAGGAAGCTGTGTATATACAAAGGAATGCTGAAATGTGTTG	2196
QY	3094	GTCTTCAATGTCTCAAAAGTAAATGGGGGATCGGCTACCGCTGAGCATGTACATAG	3153
Db	2197	GTCTTCAATGTCTCAAAAGTAAATGGGGGATCGGCTACCGCTGAGCATGTACATAG	2256
QY	3154	ACAAATATGTGCCACAGAAATCTTTCTTCACTGTTTAAACACATATACCTGGAGCTA	3213
Db	2257	ACAAATATGTGCCACAGAAATCTTTCTTCACTGTTTAAACACATATACCTGGAGCTA	2316
QY	3214	CTTTATTAACAAGATGACCAACAACTGTGTATAGCTTTGCCCTTCAAGGACATGAGCA	3273
Db	2317	CTTTATTAACAAGATGACCAACAACTGTGTATAGCTTTGCCCTTCAAGGACATGAGCA	2376
QY	3274	AAATTTGAGTTGTTTCTGCCCTAGACATCATTCAAATTTGGGTGTCATTTCTTATG	3333
Db	2377	AAATTTGAGTTGTTTCTGCCCTAGACATCATTCAAATTTGGGTGTCATTTCTTATG	2436
QY	3334	GTGTTTCCATGACACTTTGGAAGACCTATTTTAAAGCTAGAGTTGGAAGCAAAATTC	3393
Db	2437	GTGTTTCCATGACACTTTGGAAGACCTATTTTAAAGCTAGAGTTGGAAGCAAAATTC	2496
QY	3394	ACCAAGCAGATATTAGTGTATTTACTCAGACGCCACTGGAGGAGAAATGGATTCAAAAT	3453
Db	2497	ACCAAGCAGATATTAGTGTATTTACTCAGACGCCACTGGAGGAGAAATGGATTCAAAAT	2556
QY	3454	CTTTTGTATGAATGGAAACAGAGCTTACTTATCTTCTGAAACCAAGGCTTCTTAGTGA	3513
Db	2557	CTTTTGTATGAATGGAAACAGAGCTTACTTATCTTCTGAAACCAAGGCTTCTTAGTGA	2616
QY	3514	GCACCATGAGCTTTGGAACCAACAGATGTATCAATAGCAAGTTTCATTTCTTACCT	3573
Db	2617	GCACCATGAGCTTTGGAACCAACAGATGTATCAATAGCAAGTTTCATTTCTTACCT	2676
QY	3574	TGAAACCTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG	3633

Db	2677	TGAAACCTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG	2736
QY	3634	TTCCAGATTTTATGTTTGTGTTTTCATCACTCTTTTAAATATGCTGTTCCCATCAAAC	3693
Db	2737	TTCCAGATTTTATGTTTGTGTTTTCATCACTCTTTTAAATATGCTGTTTCCCATCAAAC	2796
QY	3694	TTGTTCCAGACTTATATTTTCTAAAACTGGAGACAAACACATAAATACAAACAAAGTC	3753
Db	2797	TTGTTCCAGACTTATATTTTCTAAAACTGGAGACAAACACATAAATACAAACAAAGTC	2856
QY	3754	TGCTTCTTCAAAATTTCTGCT-----GACTCAGATATCAGTGATCTTA	3795
Db	2857	TGCTTCTTCAAAATTTCTGCTGTGAGAGTGTNNGTGAAGACTCAGATATCAGTGATCTTA	2916
QY	3796	TTAGCTTTTTCACAAGCCAGAACATAATGCTGACGATGATTAATGACAGTGACTATGTAT	3855
Db	2917	TTAGCTTTTTCACAAGCCAGAACATAATGCTGACGATGATTAATGACAGTGACTATGTAT	2976
QY	3856	CGTGGCTCCCATAGTGGGGCTTTAAATGTGATGCAITTCAGAAAAAGGACTATGTTTTTG	3915
Db	2977	CGTGGCTCCCATAGTGGGGCTTTAAATGTGATGCAITTCAGAAAAAGGACTATGTTTTTG	3036
QY	3916	CAGCTGTTTTCACAAGTACTATGTTTATCTTTTAAATGTGATGCAITTCAGTTAGTAT	3975
Db	3037	CAGCTGTTTTCACAAGTACTATGTTTATCTTTTAAATGTGATGCAITTCAGTTAGTAT	3096
QY	3976	ACTACTATCTTTTAAATGTGATGCAITTCAGAACCATCCAGATCTCGAGTACCCCATCT	4035
Db	3097	ACTACTATCTTTTAAATGTGATGCAITTCAGAACCATCCAGATCTCGAGTACCCCATCT	3156
QY	4036	TTCAAGAAATTTACTGATATAGTTTTTAAATGTGAGCTGTATTTTCAAGCAGCTTTGCTTG	4095
Db	3157	TTCAAGAAATTTACTGATATAGTTTTTAAATGTGAGCTGTATTTTCAAGCAGCTTTGCTTG	3216
QY	4096	G-AATCATTTGTTACTGCAATGCCACTTACTTTGCCATGGAAATGCAGAGATCATAG	4154
Db	3217	GAAATCATTTGTTAACTGCAATGCCACTTACTTTGCCATGGAAATGCAGAGATCATATA	3274
QY	4155	ATCAAGCTTATACTCAACTTAACTTTTCAAGCTTCTTCCCATCTGCATATTTGGATGGA	4214
Db	3275	-----AGGCTTTTGGCCATCTGCATATTTGGATTTGGA	3305
QY	4215	CAAGCTGTTGTTGATATCCCTTATTTTATCATTTTATTTGATGCTAGGAAGCTTA	4274
Db	3306	CAAGCTGTTGTTGATATCCCTTATTTTATCATTTTATTTGATGCTAGGAAGCTTA	3365
QY	4275	CTGGCAATTTCAATATGGAATTAATTTTATACGTAAAGTTCTGCTGTGTTTTTGC	4334
Db	3366	TTGGCAATTTCAATATGGAATTAATTTTATACGTAAAGTTCTGCTGTGTTTTTGC	3425
QY	4335	CTTATTTGTTATGTTCCATCAGTTATCTGTTCACTTATATTTGCTTTTTCACCTTTAG	4394
Db	3426	CTTATTTGTTATGTTCCATCAGTTATCTGTTCACTTATATTTGCTTTTTCACCTTTAG	3485
QY	4395	AAAAATTTTAAATACCAAGAAATTTTGGTCAATTTATCTATTTCTGTGGCAGGTTGCTTGT	4454
Db	3486	AAAAATTTTAAATACCAAGAAATTTTGGTCAATTTATCTATTTCTGTGGCAGGTTGCTTGT	3545
QY	4455	ATTGCAATCACTGAAATAAATTTCTTTATGGGATACAAATTTGCAACTATTTCTTCAATAT	4514
Db	3546	ATTGCAATCACTGAAATAAATTTCTTTATGGGATACAAATTTGCAACTATTTCTTCAATAT	3605
QY	4515	GCCTTTTGTATGTTCCATCAGTTATCCATCTATCCATCTAGGTTGCCGTGATTTCTTCAATAG	4574
Db	3606	GCCTTTTGTATGTTCCATCAGTTATCCATCTATCCATCTAGGTTGCCGTGATTTCTTCAATAG	3665
QY	4575	ATTTCTTGGAGAAATGTCAGAAAAATGTGGACACCTATAATCCATGGGATAGGCTTTCA	4634
Db	3666	ATTTCTTGGAGAAATGTCAGAAAAATGTGGACACCTATAATCCATGGGATAGGCTTTCA	3725
QY	4635	GTAGCTGTTTATATCGCTTACCTGCGAGTGTACTGTGGATTTTCTTCTTACAAATACTAT	4694



therapeutic, diagnostic and pharmacogenomic applications -

Disclosure; Page 45-46; 46pp; English.

The present sequence is a cDNA coding for novel human protein (NHP), human transporter protein. NHPs shares structural similarity with the mammalian ATP-binding cassette (ABC) transporters and multidrug resistance transporters. NHP polynucleotides are useful for the therapeutic, diagnostic and pharmacogenomic applications. They are used for detecting and treating mental disorders and cancers. They are also used in gene therapy. NHP polypeptides are useful for diagnosis, drug screening, clinical trial monitoring, treatment of diseases and disorders, and cosmetic or nutraceutical applications.

Sequence 5262 BP; 1615 A; 888 C; 1012 G; 1743 T; 4 other;

Query Match 80.48; Score 5247.4; DB 24; Length 5262;  
Best Local Similarity 99.88; Pred. No. 0;  
Matches 5248; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy	757	ACTGTTGATATGGTGGTATTTCAAATTCGTCTACCCCTATTTTCACATGCTTGTCTTACT	816
Db	1	ACTGTTGATATGGTGGTATTTCAAATTCGTCTACCCCTATTTTCACATGCTTGTCTTACT	60
Qy	817	TTTCAGAGCTGACAGATTGCTGCTCCATGCAATCTGCTCCAGTTTCTTAAGAGACAGCT	876
Db	61	TTTCAGAGCTGACAGATTGCTGCTCCATGCAATCTGCTCCAGTTTCTTAAGAGACAGCT	120
Qy	877	TGGAGTATGCTTAATCATCTTACCTGGGACTGAAACAGCTGCTTATTTTCCGTTAAAA	936
Db	121	TGGAGTATGCTTAATCATCTTACCTGGGACTGAAACAGCTGCTTATTTTCCGCTTAAAA	180
Qy	937	ATTACATGCAAGTTTACCTGCGTGGCTCCGGTGTGTTGTTGTTTCTCTTAAATAGG	996
Db	181	ATTACATGCAAGTTTACCTGCGTGGCTCCGGTGTGTTGTTGTTTCTCTTAAATAGG	240
Qy	997	TTTATTCCAGAAAAATGCTCACTGCAATTAGGGAGGTAGGAGTTTGGAGACAGACAGAA	1056
Db	241	TTTATTCCAGAAAAATGCTCACTGCAATTAGGGAGGTAGGAGTTTGGAGACAGACAGAA	300
Qy	1057	CACCTTCTACTGAAGAATTACTTAATTAATGCAGAACCAAAAGAGTAGTGTTCAGGAAA	1116
Db	301	CACCTTCTACTGAAGAATTACTTAATTAATGCAGAACCAAAAGAGTAGTGTTCAGGAAA	360
Qy	1117	TTCTTTTCCACTATTTTATTTTGGTTTAAATTAATTAATGATCATGCTCATCCCAATA	1176
Db	361	TTCTTTTCCACTATTTTATTTTGGTTTAAATTAATTAATGATCATGCTCATCCCAATA	420
Qy	1177	AGAAATATGAAGAGTCCCTAATATAGAACTCAATCCTATGGCAAGTTTACTCTTTCTA	1236
Db	421	AGAAATATGAAGAGTCCCTAATATAGAACTCAATCCTATGGCAAGTTTACTCTTTCTA	480
Qy	1237	ATCTAATTTCTGGATATCTCCAGTGACTAATTAATTAATTAATGATCATGCTCATGAAAGTGT	1296
Db	481	ATCTAATTTCTGGATATCTCCAGTGACTAATTAATTAATTAATGATCATGCTCATGAAAGTGT	540
Qy	1297	CTACTGATCATCTACCTGATGCTCAATAATTAATGAAGAATATACAAATGAAAGAAATGT	1356
Db	541	CTACTGATCATCTACCTGATGCTCAATAATTAATGAAGAATATACAAATGAAAGAAATGT	600
Qy	1357	TAAACATCCAGTCTCTTAAGCCGAGCAACTTTGTAGGTGGTTTCCAAAGACTCCATGT	1416
Db	601	TAAACATCCAGTCTCTTAAGCCGAGCAACTTTGTAGGTGGTTTCCAAAGACTCCATGT	660
Qy	1417	CCTATGAACCTTGGTTTTTCTCGATATGATTCAGATATCTTCTATTTATATGATGATTCAA	1476
Db	661	CCTATGAACCTTGGTTTTTCTCGATATGATTCAGATATCTTCTATTTATATGATGATTCAA	720
Qy	1477	GAGCTGGCTGTTCAAAATCATGTGAGCTGCTCAGTACTGCTCTCAGGTTTCCACAGTTT	1536
Db	721	GAGCTGGCTGTTCAAAATCATGTGAGCTGCTCAGTACTGCTCTCAGGTTTCCACAGTTT	780
Qy	1537	TACAAGCATCCATAGATGCTGCCATTATACAGTTTGAAGACCAATGTTTCTCTTTTGAAGG	1596

Db	781	TACAAGCATCCATAGATGCTGCCATTATATACAGTTGAAGACCAATGTTTCTCTTTGAAGG	840
Qy	1597	AGCTGAGAGCAACTAAAGCTGTTTATTTATGGAGAAACTGCTGTGTAGAAATAGATACCT	1656
Db	841	AGCTGAGAGCAACTAAAGCTGTTTATTTATGGAGAAACTGCTGTGTAGAAATAGATACCT	900
Qy	1657	TTCCCGAGAGAGTAAATTTTAAATATACCTAGTTATAGCAATTTTCCACCTTTTGGTACTTTT	1716
Db	901	TTCCCGAGAGAGTAAATTTTAAATATACCTAGTTATAGCAATTTTCCACCTTTTGGTACTTTT	960
Qy	1717	TGCAATTTATATCTGATAGAGAAAAAGAAAAATTAAGAAATTTTAAAGATATAGG	1776
Db	961	TGCAATTTATATCTGATAGAGAAAAAGAAAAATTAAGAAATTTTAAAGATATAGG	1020
Qy	1777	GACTTCATGATGCTGCTTTTGGCTTTTCTGGGTCTCTCTATATACAAAGTTTAAATTTTTC	1836
Db	1021	GACTTCATGATGCTGCTTTTGGCTTTTCTGGGTCTCTCTATATACAAAGTTTAAATTTTTC	1080
Qy	1837	TTATGTCCCTTCTTATGGCAGTCAATGGCAGCACTTCTTTTGTATTTTCTCAAAGTAGCA	1896
Db	1081	TTATGTCCCTTCTTATGGCAGTCAATGGCAGCACTTCTTTTGTATTTTCTCAAAGTAGCA	1140
Qy	1897	GCATTTGATATTTTCTGCTTTTCTTATGGATTAATCATCTGTATTTTGTCTTTAA	1956
Db	1141	GCATTTGATATTTTCTGCTTTTCTTATGGATTAATCATCTGTATTTTGTCTTTAA	1200
Qy	1957	TGCTGACACTCTTTTAAATAATCAAAACATGCTGGGAATAGTTGAATTTTGTGTACTG	2016
Db	1201	TGCTGACACTCTTTTAAATAATCAAAACATGCTGGGAATAGTTGAATTTTGTGTACTG	1260
Qy	2017	TGCTTTTGGATTTATTTGGCTTTATGATTAATCTCATAGAAAGTTTCCCAATCGTTAG	2076
Db	1261	TGCTTTTGGATTTATTTGGCTTTATGATTAATCTCATAGAAAGTTTCCCAATCGTTAG	1320
Qy	2077	TGTGGCTTTTCACTCTTCTGTCACTGTTTGTGATTTGTGATTTGCAAGGTCAAGC	2136
Db	1321	TGTGGCTTTTCACTCTTCTGTCACTGTTTGTGATTTGTGATTTGCAAGGTCAAGC	1380
Qy	2137	ATTAGAAGATTTTAAATGAAGTGTCTTCAATTTTGAAGTGTGCTGAGGCCCAATCCCTC	2196
Db	1381	ATTAGAAGATTTTAAATGAAGTGTCTTCAATTTTGAAGTGTGCTGAGGCCCAATCCCTC	1440
Qy	2197	TAATTTATCAATTTATCATGCTCACACTTAAATAGTATATTTCTATGCTCTTGGCTGTCT	2256
Db	1441	TAATTTATCAATTTATCATGCTCACACTTAAATAGTATATTTCTATGCTCTTGGCTGTCT	1500
Qy	2257	ATCTTGATCAAGTCAATTCAGGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGA	2316
Db	1501	ATCTTGATCAAGTCAATTCAGGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGA	1560
Qy	2317	AGCTTTATATTTGGTCAAGAGTAAAGAAATTTATGAGGATTTATCAGAGGCAATGTTA	2376
Db	1561	AGCTTTATATTTGGTCAAGAGTAAAGAAATTTATGAGGATTTATCAGAGGCAATGTTA	1620
Qy	2377	ATGGAATATTTAGTTTGTAGTAAATTTTGAAGCAGTTTCTTCAAGAAATTTGTAGGAAAAG	2436
Db	1621	ATGGAATATTTAGTTTGTAGTAAATTTTGAAGCAGTTTCTTCAAGAAATTTGTAGGAAAAG	1680
Qy	2437	AAGCCATAAGAAATTTAGTGGTATTTCAAGAGACATACAGAAAGAGGGTGAAGATGAGG	2496
Db	1681	AAGCCATAAGAAATTTAGTGGTATTTCAAGAGACATACAGAAAGAGGGTGAAGATGAGG	1740
Qy	2497	CTTTGAGAAATTTGCTATTTGACATATATGAGGTCAGATTAAGTCTTCTTCTTGGCCACA	2556
Db	1741	CTTTGAGAAATTTGCTATTTGACATATATGAGGTCAGATTAAGTCTTCTTCTTGGCCACA	1800
Qy	2557	GTGGAACAGGAAAGAGTACATTTGATCAATTTCTTTTGTGAGCTCTGCCACCTTCTGATG	2616
Db	1801	GTGGAACAGGAAAGAGTACATTTGATCAATTTCTTTTGTGAGCTCTGCCACCTTCTGATG	1860
Qy	2617	GGTTTGCATCTATATATGACACAGAGTCTCAGAAATAGATGAATTTTGTGAAGCAAGAA	2676

1861 GGTTTGCACTATATATGACACAGAGCTCTCAGAAATAGATGAAATGTTTGAAGCAAGAA 1920  
Qy  
2677 AAATGATGCAATTTGTCACAGCTAGATATACACTTTGATGTTTGAAGTAGAGAA 2736  
Db  
1921 AAATGATGTCATTTGTCACAGCTAGATATACACTTTGATGTTTGAAGTAGAGAA 1980  
Db  
2737 ATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAACATATAATACAAAGAGTGC 2796  
Qy  
1981 ATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAACATATAATACAAAGAGTGC 2040  
Db  
2797 AGAAGGTTTTACTAGATTTAGACATGACAGACTATCAAGATACCAAGCTTAAATAATTA 2856  
Qy  
2041 AGAAGGTTTTACTAGATTTAGACATGACAGACTATCAAGATACCAAGCTTAAATAATTA 2100  
Db  
2857 GTGGTGTCAAAAAGAAAGCTGTCTATTAGGAATGCTGTTCTTGGGAACCCAAAGATAC 2916  
Qy  
2101 GTGGTGTCAAAAAGAAAGCTGTCTATTAGGAATGCTGTTCTTGGGAACCCAAAGATAC 2160  
Db  
2917 TGCTGCTAGATGAACCAAGCTGGAAATGACCCCTGTTCTCGACATATTTGATGGAATC 2976  
Qy  
2161 TGCTGCTAGATGAACCAAGCTGGAAATGACCCCTGTTCTCGACATATTTGATGGAATC 2220  
Db  
2977 TTTTAAATACAGAAAGCCAAATCGGGTGAAGCTGTTCACTACTCAATTTCAATGATGAAG 3036  
Qy  
2221 TTTTAAATACAGAAAGCCAAATCGGGTGAAGCTGTTCACTACTCAATTTCAATGATGAAG 2280  
Db  
3037 CTGACATTTCTGCAGATAGAAAGCTGTGATATACAAAGAAATGCTGAAATGTTGGTT 3096  
Qy  
2281 CTGACATTTCTGCAGATAGAAAGCTGTGATATACAAAGAAATGCTGAAATGTTGGTT 2340  
Db  
3097 CTTCAATGTTTCTTCAAAAGTAAATGGGGATCGGCTACCCCTGAGCATGTACATAGACA 3156  
Qy  
2341 CTTCAATGTTTCTTCAAAAGTAAATGGGGATCGGCTACCCCTGAGCATGTACATAGACA 2400  
Db  
3157 AATATTGTCGCCACAGAAATCTCTTTCTCACTGGTTAAACAAACATATACCTGGAGCTACTT 3216  
Qy  
2401 AATATTGTCGCCACAGAAATCTCTTTCTCACTGGTTAAACAAACATATACCTGGAGCTACTT 2460  
Db  
3217 TATTACAAAGATGACCAACACTTGTGTATAGCTTGGCTTTCAAGGACATGGCAAAAT 3276  
Qy  
2461 TATTACAAAGATGACCAACACTTGTGTATAGCTTGGCTTTCAAGGACATGGCAAAAT 2520  
Db  
3277 TTTTCAGGTTTGTCTTGGCTTACAGAGCTCAATTTCAAAATTTGGGTGTCAATTTCTATGGTG 3336  
Qy  
2521 TTTTCAGGTTTGTCTTGGCTTACAGAGCTCAATTTCAAAATTTGGGTGTCAATTTCTATGGGG 2580  
Db  
3337 TTTTCATGACGACTTTGGGAAGAGCTATTTTTAAAGCTTAGAAGTTGAAGCAGAAATGACC 3396  
Qy  
2581 TTTTCATGACGACTTTGGGAAGAGCTATTTTTAAAGCTTAGAAGTTGAAGCAGAAATGACC 2640  
Db  
3397 AAGCAGATTTATAGTGTATTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAATCTT 3456  
Qy  
2641 AAGCAGATTTATAGTGTATTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAATCTT 2700  
Db  
3457 TTGATGAATTTGGAACACAGAGCTTACTTATCTTTCTGAAACCAAGGCTTCTAGTAGGACA 3516  
Qy  
2701 TTGATGAATTTGGAACACAGAGCTTACTTATCTTTCTGAAACCAAGGCTTCTAGTAGGACA 2760  
Db  
3517 CCATGAGCCTTTTGGAAACAAACAGATGTATCAATAGCAAAAGTTTCAATTTCTTTACCTTGA 3576  
Qy  
2761 CCATGAGCCTTTTGGAAACAAACAGATGTATCAATAGCAAAAGTTTCAATTTCTTTACCTTGA 2820  
Db  
3577 AACGTGAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTC 3636  
Qy  
2821 AACGTGAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTC 2880  
Db  
3637 AGATTTTTATGTTTTTGGTTTCACTCCTTTTAAATGCTGTGGTTCCCATCAAACTTG 3696  
Qy  
2881 AGATTTTTATGTTTTTGGTTTCACTCCTTTTAAATGCTGTGGTTCCCATCAAACTTG 2940  
Db  
3697 TTCCAGACTTATATTTTCTAAAACCTGGAGACAAACACATAAATACAAAACAGTCTGC 3756  
Qy  
2941 TTCCAGACTTATATTTTCTAAAACCTGGAGACAAACACATAAATACAAAACAGTCTGC 3000  
Db

3757 TTCTTCAAAATTTCTGCTGACTCAGATATCAGTGATCTTATTAGCTTTTTTCAAGCCAGA 3816  
Qy  
3001 TTCTTCAAAATTTCTGCTGACTCAGATATCAGTGATCTTATTAGCTTTTTTCAAGCCAGA 3060  
Db  
3817 ACATAATCGTGACGATGATTAATGACAGTGACTATGATATCGTGGCTCCCATAGTGGG 3876  
Qy  
3061 ACATAATCGTGACGATGATTAATGACAGTGACTATGATATCGTGGCTCCCATAGTGGG 3120  
Db  
3877 CTTTAAATGTGATGCAATTCAGAAAGGACTATGTTTTGACAGCTGTTTTCAACGACTCTA 3936  
Qy  
3121 CTTTAAATGTGATGCAATTCAGAAAGGACTATGTTTTGACAGCTGTTTTCAACGACTCTA 3180  
Db  
3937 TGGTTTATTTCTTACTATATATTAGTGAATATCATTTAGTAACTACTATCTTTTATCATTTAA 3996  
Qy  
3181 TGGTTTATTTCTTACTATATATTAGTGAATATCATTTAGTAACTACTATCTTTTATCATTTAA 3240  
Db  
3997 ATGTGACTGAAACCAATCCAGATCTGGAGTACCCCAATCTTTCAAGAAATTAATGATATAG 4056  
Qy  
3241 ATGTGACTGAAACCAATCCAGATCTGGAGTACCCCAATCTTTCAAGAAATTAATGATATAG 3300  
Db  
4057 TTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGC 4116  
Qy  
3301 TTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGC 3360  
Db  
4117 CACCTTTACTTTGCCATGGAATAATGCAGAGAAATCATTAAGATCAAAAGCTTTATATCAACTTA 4176  
Qy  
3361 CACCTTTACTTTGCCATGGAATAATGCAGAGAAATCATTAAGATCAAAAGCTTTATATCAACTTA 3420  
Db  
4177 AACTTTCAAGGCTTTTGGCATCTGCAATATGGATGGACAAAGCTGTTGTTGATATCCCT 4236  
Qy  
3421 AACTTTCAAGGCTTTTGGCATCTGCAATATGGATGGACAAAGCTGTTGTTGATATCCCT 3480  
Db  
4237 TATTTTTTATCATTTCTTATTTTATGCTAGAGCTTACTGCAATTTTCAATTTATGATAT 4296  
Qy  
3481 TATTTTTTATCATTTCTTATTTTATGCTAGAGCTTATTTGCAATTTTCAATTTATGATAT 3540  
Db  
4297 ATTTTTTATCTGTAAGTTTCTTGTGCTGTTTGTGCTTTATTTGCTTATTTGCTTATGCTCATCAG 4356  
Qy  
3541 ATTTTTTATCTGTAAGTTTCTTGTGCTGTTTGTGCTTTTATTTGCTTATTTGCTTATGCTCATCAG 3600  
Db  
4357 TTATTTCTGTTCACTTATATGCTTTCTTCACTTTTAAAGAAATTTTAAATACCAAGAAAT 4416  
Qy  
3601 TTATTTCTGTTCACTTATATGCTTTCTTCACTTTTAAAGAAATTTTAAATACCAAGAAAT 3660  
Db  
4417 TTTGTCATTTTATCTATTTCTGTCGACGCTTGTGCTTATTTGCAATCTACTGAAATAACTT 4476  
Qy  
3661 TTTGTCATTTTATCTATTTCTGTCGACGCTTGTGCTTATTTGCAATCTACTGAAATAACTT 3720  
Db  
4477 TCTTTTATGGGATACAAATTTGCAACTTATTTCTTATGCTTTTGTATCATCTTCCAA 4536  
Qy  
3721 TCTTTTATGGGATACAAATTTGCAACTTATTTCTTATGCTTTTGTATCATCTTCCAA 3780  
Db  
4537 TCTATCCACTTTCTAGGTTGCTGATTTCTTTTATATAAGATTTCTTTGGAAGAAATGACGAA 4596  
Qy  
3781 TCTATCCACTTTCTAGGTTGCTGATTTCTTTTATATAAGATTTCTTTGGAAGAAATGACGAA 3840  
Db  
4597 AAAATGTGACACCTTATTAATCCATGGGATAGGCTTTGATAGCTGTTTATATCGCTTACC 4656  
Qy  
3841 AAAATGTGACACCTTATTAATCCATGGGATAGGCTTTGATAGCTGTTTATATCGCTTACC 3900  
Db  
4657 TGCAGTGTGATCTGTGGATTTTCTTCTTCAATCTATGAGAAATAATATGAGGCGAGAT 4716  
Qy  
3901 TGCAGTGTGATCTGTGGATTTTCTTCTTCAATCTATGAGAAATAATATGAGGCGAGAT 3960  
Db  
4717 CAATAAGAAAAGATCCCTTTTTCAGAAAACCTTTTCAAGAAAGCTTAAATAATAGAAAGCTTC 4776  
Qy  
3961 CAATAAGAAAAGATCCCTTTTTCAGAAAACCTTTTCAAGAAAGCTTAAATAATAGAAAGCTTC 4020  
Db  
4777 CAGAAACCAACAGCAATCAGAGATGAAGATGAAGATGTCAAAGCTTAAAGACTTAAAGGTCA 4836  
Qy  
4021 CAGAAACCAACAGCAATCAGAGATGAAGATGAAGATGTCAAAGCTTAAAGACTTAAAGGTCA 4080  
Db

QY 4837 AAGAGCTGATGGGTGCGAGTGTGTGAGGAGAAACCATCATATATGCTCAGCAATTTGC 4896  
DB 4081 AAGAGCTGATGGGTGCGAGTGTGTGAGGAGAAACCATCATATATGCTCAGCAATTTGC 4140  
QY 4897 ATAAGAATATGATGACAAGAGATTTCTCTTTCAAGAAAGTAAAGAAAGTGGCAA 4956  
DB 4141 ATAAGAATATGATGACAAGAGATTTCTCTTTCAAGAAAGTAAAGAAAGTGGCAA 4200  
QY 4957 CTAATATCATCTCTTTCTGTGTGAAAAAGAGAGATCTTTAGGACTATTTGGTCCAAATG 5016  
DB 4201 CTAATATCATCTCTTTCTGTGTGAAAAAGAGAGATCTTTAGGACTATTTGGTCCAAATG 4260  
QY 5017 GTGCTGCAAAAGACAAATTAATTAATTTCTGTTGTGTGATATTAAGAAAGTAAAGTGGCAA 5076  
DB 4261 GTGCTGCAAAAGACAAATTAATTAATTTCTGTTGTGTGATATTAAGAAAGTAAAGTGGCAA 4320  
QY 5077 AGGTATTTTATGAGATTAATTTCTCAGAGCAAGTGAAGATGATGATCTCACTGAAGTGA 5136  
DB 4321 AGGTATTTTATGAGATTAATTTCTCAGAGCAAGTGAAGATGATGATCTCACTGAAGTGA 4380  
QY 5137 TGGGTTACTGCTCCTCAGATAAACCCCTTTGTGGCCAGATACATCTCAGGAAACATTTTG 5196  
DB 4381 TGGGTTACTGCTCCTCAGATAAACCCCTTTGTGGCCAGATACATCTCAGGAAACATTTTG 4440  
QY 5197 AAATTTATGAGCTGTCAAGGAATGATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 5256  
DB 4441 AAATTTATGAGCTGTCAAGGAATGATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4500  
QY 5257 TAACACATGACTGATTAAGAAAGATCTCTCAGAGACATGTAAGAAAGTGAAGTGAAGTGA 5316  
DB 4501 TAACACATGACTGATTAAGAAAGATCTCTCAGAGACATGTAAGAAAGTGAAGTGAAGTGA 4560  
QY 5317 GAATCAACGAAGTGTGTTTGTCTCTAAGTATGCTAGGAAATCTCTCAGATTAATTTGC 5376  
DB 4561 GAATCAACGAAGTGTGTTTGTCTCTAAGTATGCTAGGAAATCTCTCAGATTAATTTGC 4620  
QY 5377 TAGATGAACCATCTACAGGTATGATCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAATTC 5436  
DB 4621 TAGATGAACCATCTACAGGTATGATCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAATTC 4680  
QY 5437 GAATGCAATTTAAAGCAAGAGCGGCTGCTATCTGACCACTCACTATATGAGGAGG 5496  
DB 4681 GAATGCAATTTAAAGCAAGAGCGGCTGCTATCTGACCACTCACTATATGAGGAGG 4740  
QY 5497 CAGAGGCTGTCTGATCGAGTAGCTATCATGTTGCTGTGGCAGTTAAGATGATCGAA 5556  
DB 4741 CAGAGGCTGTCTGATCGAGTAGCTATCATGTTGCTGTGGCAGTTAAGATGATCGAA 4800  
QY 5557 CAGTACCAATCTAAAGAGTAAATTTTGGAAAGGCTACTTTTGGAAATTTAAATTTGAAG 5616  
DB 4801 CAGTACCAATCTAAAGAGTAAATTTTGGAAAGGCTACTTTTGGAAATTTAAATTTGAAG 4860  
QY 5617 ACTGGATAGAAACCTTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAA 5676  
DB 4861 ACTGGATAGAAACCTTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAA 4920  
QY 5677 ATGCAAGCCGTGAGAAAGTGTCTCTATTTTGGCTTATAAAATTCCTAAGGAAGATG 5736  
DB 4921 ATGCAAGCCGTGAGAAAGTGTCTCTATTTTGGCTTATAAAATTCCTAAGGAAGATG 4980  
QY 5737 TTCAGTCCCTTTCACAACTCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGGCAATG 5796  
DB 4981 TTCAGTCCCTTTCACAACTCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGGCAATG 5040  
QY 5797 AAGATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAGAAC 5856  
DB 5041 AAGATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAGAAC 5100  
QY 5857 AAGAGGAGGAGTAATAGTTGTGGAATCTTTAAACAGCACATTTTGTGGGAGGACAC 5916  
DB 5101 AAGAGGAGGAGTAATAGTTGTGGAATCTTTAAACAGCACATTTTGTGGGAGGACAC 5160  
QY 5917 AAGAGATAGAGTAGTATTTTGAATTTGTATTTGTTGCTGCTTACTGGGACTTCTTTC 5976

DB 5161 AAGAAGATAGTAGTATTTTGAATTTGATTTGCTGCTGCTTACTGGGACTTCTTTC 5220  
QY 5977 TTTTTCACCTTAATTTTAACTTTGTTGTTTAAAAAGTTTTT 6014  
DB 5221 TTTTTCACCTTAATTTTAACTTTGTTGTTTAAAAAGTTTTT 5258  
RESULT 6  
AAD37618  
ID AAD37618 standard; cDNA; 4929 BP.  
XX  
AC AAD37618;  
XX  
DT 10-SEP-2002 (first entry)  
XX  
DE Human transporter protein cDNA #1.  
XX  
KW Human; novel human protein; NHP; transporter protein; mental disorder;  
KW cancer; gene therapy; drug screening; nutraceutical application; gene;  
KW cosmetic application; polymorphism; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 1..4929  
FT /\*tag= a  
FT /product= "Human transporter protein #1"  
FT /transl\_except= (pos:808..810, aa:Tyr)  
FT /transl\_except= (pos:2494..2496, aa:Ser)  
FT /transl\_except= (pos:2878..2880, aa:Val)  
FT misc\_feature 810  
FT /\*tag= b  
FT /note= "This degenerate base represents a  
FT polymorphic site"  
FT misc\_feature 2494  
FT /\*tag= c  
FT /note= "This degenerate base represents a  
FT polymorphic site"  
FT misc\_feature 2878  
FT /\*tag= d  
FT /note= "This degenerate base represents a  
FT polymorphic site"  
XX  
WO200231147-A2.  
XX  
PD 18-APR-2002.  
XX  
PF 04-OCT-2001; 2001WO-US311113.  
XX  
PR 10-OCT-2000; 2000US-239629P.  
XX  
PA (LEXI-) LEXICON GENETICS INC.  
XX  
PI Hu Y, Nepomnichy B;  
XX  
DR WPI: 2002-452347/48.  
DR P-PSDB; AAE23656.  
XX  
PT Novel nucleic acid molecule encoding novel human proteins, useful for  
PT therapeutic, diagnostic and pharmacogenomic applications  
XX  
PS Claim 1; Page 34-35; 46pp; English.  
XX  
CC The present sequence is a cDNA coding for novel human protein (NHP),  
CC human transporter protein. NHPs shares structural similarity with  
CC the mammalian ATP-binding cassette (ABC) transporters and multidrug  
CC resistance transporters. NHP polynucleotides are useful for the  
CC therapeutic, diagnostic and pharmacogenomic applications. They are  
CC used for detecting and treating mental disorders and cancers. They  
CC are also used in gene therapy. NHP polypeptides are useful for  
CC diagnosis, drug screening, clinical trial monitoring, treatment of  
CC diseases and disorders, and cosmetic or nutraceutical applications.

XX	Sequence	4929 BP; 1544 A; 827 C; 950 G; 1604 T; 4 other;	
SQ	Query Match	75.4%; Score 4918.4; DB 24; Length 4929;	
	Best Local Similarity	99.8%; Pred. No. 0;	
	Matches 4919; Conservative	4; Mismatches 6; Indels 0; Gaps 0;	
QY	1011	ATGTCCTCACTCAATTTAGGGAGGTAGGAGTTTGGAGACAGACCAAGCACTTCTACTGAAG	1070
DB	1	ATGTCCTCACTGCAATTTAGGGAGGTAGGAGTTTGGAGACAGACCAAGCACTTCTACTGAAG	60
QY	1071	AATTACTTAATTAATGCAGAACCAAAAAGAGTAGTGTTCAGGAAATTTCTTTTCCACTA	1130
DB	61	AATTACTTAATTAATGCAGAACCAAAAAGAGTAGTGTTCAGGAAATTTCTTTTCCACTA	120
QY	1131	TTTTTTTTTATTTGGTTAAATTAATTAGCATGATCCAAATAAGAAATATGAAGAA	1190
DB	121	TTTTTTTTTATTTGGTTAAATTAATTAGCATGATCCAAATAAGAAATATGAAGAA	180
QY	1191	GTGCTAATATAGAACTCAATCTATGCACAAGTTTACTCTTTCTAATCTAAATTTCTTGGA	1250
DB	181	GTGCTAATATAGAACTCAATCTATGCACAAGTTTACTCTTTCTAATCTAAATTTCTTGGA	240
QY	1251	TATACTCCAGTGACTAATATTTACAAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTA	1310
DB	241	TATACTCCAGTGACTAATATTTACAAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTA	300
QY	1311	CTGATGCTAATATTACTGAAGAAATATACAAATGAAAGAAATGTTAAAGCATCAGTCTC	1370
DB	301	CTGATGCTAATATTACTGAAGAAATATACAAATGAAAGAAATGTTAAAGCATCAGTCTC	360
QY	1371	TCTAAGCCGAGCACTTTTGTAGGTGTGTTTCAAGACTCCATGCTCTATGAATTCGT	1430
DB	361	TCTAAGCCGAGCACTTTTGTAGGTGTGTTTCAAGACTCCATGCTCTATGAATTCGT	420
QY	1431	TTTTTCTCGATATGATTCAGTATCTTCTATTATATGATTCAGAGCTGGCTGTCTA	1490
DB	421	TTTTTCTCGATATGATTCAGTATCTTCTATTATATGATTCAGAGCTGGCTGTCTA	480
QY	1491	AAATCATGTGAGGCTGCTCAGTACTGTGCTCCTCAGTTCACAGTTTACAAGCATCGATA	1550
DB	481	AAATCATGTGAGGCTGCTCAGTACTGTGCTCCTCAGTTCACAGTTTACAAGCATCGATA	540
QY	1551	GATCTGCCATTATACAGTTTGAAGCAATGTTTCTCTTTGGAAGGAGCTGGAGTCAACT	1610
DB	541	GATCTGCCATTATACAGTTTGAAGCAATGTTTCTCTTTGGAAGGAGCTGGAGTCAACT	600
QY	1611	AAAGCTGTTATTTAGGGAGAACTGCTGTTGTAGAAATAGATACCTTTTCCCGAGGAGTA	1670
DB	601	AAAGCTGTTATTTAGGGAGAACTGCTGTTGTAGAAATAGATACCTTTTCCCGAGGAGTA	660
QY	1671	ATTTTAATATACCTAGTTATAGCATTTTACCTTTTGGATACCTTTTGGCAATTCATATC	1730
DB	661	ATTTTAATATACCTAGTTATAGCATTTTACCTTTTGGATACCTTTTGGCAATTCATATC	720
QY	1731	GTACAGAAAGAGAAATAAAGAAATTTTAAAGATAATGGCACTTCATGATACT	1790
DB	721	GTACAGAAAGAGAAATAAAGAAATTTTAAAGATAATGGCACTTCATGATACT	780
QY	1791	GCCTTTTGGCTTCTCCTGGGTTCTTCTATATACAAAGTTTAAATTTTCTATGTCCTTCTT	1850
DB	781	GCCTTTTGGCTTCTCCTGGGTTCTTCTATATACAAAGTTTAAATTTTCTATGTCCTTCTT	840
QY	1851	ATGCGAGTCAATGGACAGCTTCTTTGTTATTTCTCTCAAGTAGCAGCATTTGTATATT	1910
DB	841	ATGCGAGTCAATGGACAGCTTCTTTGTTATTTCTCTCAAGTAGCAGCATTTGTATATT	900
QY	1911	CTGCTTTTTTCTTTTATGATATCATCTGATTTTTTTTCTTAATGCTGACACCTCTT	1970
DB	901	CTGCTTTTTTCTTTTATGATATCATCTGATTTTTTTTCTTTAATGCTGACACCTCTT	960
QY	1971	TTTTAAAAAATCAAAACATGTGGGAATAGTTCAATTTTTTTTGTACTGTGGCTTTTGGATTT	2030

DB	961	TTTTAAAAAATCAAAACATGTGGGAATAGTTGAATTTTTTTGTACTGTGGCTTTTGGATTT	1020
QY	2031	ATTTGGCTTTATGATATATCTCTCATAGAAAGTTTTTCCAAATCGTTAGTGTGGCTTTTCACT	2090
DB	1021	ATTTGGCTTTATGATATATCTCTCATAGAAAGTTTTTCCAAATCGTTAGTGTGGCTTTTCACT	1080
QY	2091	CCTTTCTGTCACTGTACTTTTGTGATTTGTCACAGGTCATGCAATTTAGAGAATTTT	2150
DB	1081	CCTTTCTGTCACTGTACTTTTGTGATTTGTCACAGGTCATGCAATTTAGAGAATTTT	1140
QY	2151	AATGAAGTGTCTTCAATTTTCAATTTGACTGAGGCCCATATCTCTTAATTAATTAATTT	2210
DB	1141	AATGAAGTGTCTTCAATTTTCAATTTGACTGAGGCCCATATCTCTTAATTAATTAATTT	1200
QY	2211	ATCATGCTCACACTTAATAGTATATTTCTATGTCCTTTGGCTGTCTATCTTCATCAAGTC	2270
DB	1201	ATCATGCTCACACTTAATAGTATATTTCTATGTCCTTTGGCTGTCTATCTTCATCAAGTC	1260
QY	2271	ATTCAGGGGGAATTTGGCTTTACGGAGATCATCTTTATATTTTCTGAAGCTTTCAATTTGG	2330
DB	1261	ATTCAGGGGGAATTTGGCTTTACGGAGATCATCTTTATATTTTCTGAAGCTTTCAATTTGG	1320
QY	2331	TCAAAGAGTAAAGAAATTTATGAGGAGTTATCAGAGGGCAATGTTAATGGAATATTAGT	2390
DB	1321	TCAAAGAGTAAAGAAATTTATGAGGAGTTATCAGAGGGCAATGTTAATGGAATATTAGT	1380
QY	2391	TTTAGTGAATTTATTTGAGCCAGTTTCTCAGAAATTTGTAGGAAAGAAAGCCATAGAATTT	2450
DB	1381	TTTAGTGAATTTATTTGAGCCAGTTTCTCAGAAATTTGTAGGAAAGAAAGCCATAGAATTT	1440
QY	2451	AGTGGTATTTCAAGAACATACACAAAGAGGTTGAAATTTGTGAGGCTTTTGAGAAATTTTG	2510
DB	1441	AGTGGTATTTCAAGAACATACACAAAGAGGTTGAAATTTGTGAGGCTTTTGAGAAATTTTG	1500
QY	2511	TCATTTGACATATATGAGGTCAGATTTACTGCTTTACTTTGGCCACAGTGGAAACAGAAAG	2570
DB	1501	TCATTTGACATATATGAGGTCAGATTTACTGCTTTACTTTGGCCACAGTGGAAACAGAAAG	1560
QY	2571	AGTACATTTGATGAATTTCTTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATA	2630
DB	1561	AGTACATTTGATGAATTTCTTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATA	1620
QY	2631	TATGGACACAGAGTCTCAGAAATAGATGAAATTTGTTGAAGCAAGAAATAATGTTGCAAT	2690
DB	1621	TATGGACACAGAGTCTCAGAAATAGATGAAATTTGTTGAAGCAAGAAATAATGTTGCAAT	1680
QY	2691	TGTCCACAGTTAGATATACATTTGATGTTTGAAGTAAAGAAATAATTTTCAATTTTG	2750
DB	1681	TGTCCACAGTTAGATATACATTTGATGTTTGAAGTAAAGAAATAATTTTCAATTTTG	1740
QY	2751	GCCTTCAATCAAGGGATACCGCCCAATATAATACAAAGAGTGCAGAAAGTTTACTATA	2810
DB	1741	GCCTTCAATCAAGGGATACCGCCCAATATAATACAAAGAGTGCAGAAAGTTTACTATA	1800
QY	2811	GATTTAGACATCGACATATCAAGATAACCAAGCTTAAAGTAAAGTAAAGTAAAGTAAAGT	2870
DB	1801	GATTTAGACATCGACATATCAAGATAACCAAGCTTAAAGTAAAGTAAAGTAAAGTAAAGT	1860
QY	2871	AGAAAGCTGTCTATTAGGAATTCCTGTTTGGGAACCCAAAGATCTGCTGCTAGATGAA	2930
DB	1861	AGAAAGCTGTCTATTAGGAATTCCTGTTTGGGAACCCAAAGATCTGCTGCTAGATGAA	1920
QY	2931	CCAAAGCTGGGAATCGACCCCTGTTCTGACATATTTGATGGAATCTTTTAAATAACAGA	2990
DB	1921	CCAAAGCTGGGAATCGACCCCTGTTCTGACATATTTGATGGAATCTTTTAAATAACAGA	1980
QY	2991	AAAGCAATCGGGTGCAGTGTTCAGTACTCATTTTCATGGAATGAGCTGACATTTCTTGCA	3050
DB	1981	AAAGCAATCGGGTGCAGTGTTCAGTACTCATTTTCATGGAATGAGCTGACATTTCTTGCA	2040
QY	3051	GATAGAAAGCTGTGATATCAAGGAATGCTGAAATGTGTTGGTTCCTCAATGTCTCTC	3110
DB	2041	GATAGAAAGCTGTGATATCAAGGAATGCTGAAATGTGTTGGTTCCTCAATGTCTCTC	2100

QY	3111	AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGGCCACA	3170
DB	2101	AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGGCCACA	2160
QY	3171	GAATCTCTTTCTTCACCTGGTTAAACACATATACCTGGAGCTACTTTATTAACAAGAAAT	3230
DB	2161	GAATCTCTTTCTTCACCTGGTTAAACACATATACCTGGAGCTACTTTATTAACAAGAAAT	2220
QY	3231	GACCAACAACTTGTTGATAGCTTCCCTTCAAGGACATGACAAATTTTCAGGTGTTGTTT	3290
DB	2221	GACCAACAACTTGTTGATAGCTTCCCTTCAAGGACATGACAAATTTTCAGGTGTTGTTT	2280
QY	3291	TCTGCCCTAGACAGCTCAATCAAAATTTGGGTGTCAATTTCTTATGGTGTTCACATGACGACT	3350
DB	2281	TCTGCCCTAGACAGCTCAATCAAAATTTGGGTGGCAATTTCTTATGGGTTCACATGACGACT	2340
QY	3351	TTGGAAGACGTATTTTAAAGCTAGAAAGTTGAAGCAGAAATTTGACCAAGCAGATTAAGT	3410
DB	2341	TTGGAAGACGTATTTTAAAGCTAGAAAGTTGAAGCAGAAATTTGACCAAGCAGATTAAGT	2400
QY	3411	GTATTTACTCAGCAGCAGCTGGAGGAAGAAATGGATTCAAATCTTTTGATGAATGGAA	3470
DB	2401	GTATTTACTCAGCAGCAGCTGGAGGAAGAAATGGATTCAAATCTTTTGATGAATGGAA	2460
QY	3471	CAGAGCTTACTTATCTTTCTGAAACCAAGCTTCTCTAGTGAGCACCATGAGCCTTTGG	3530
DB	2461	CAGAGCTTACTTATCTTTCTGAAACCAAGCTTCTCTAGTGAGCACCATGAGCCTTTGG	2520
QY	3531	AAACAAACAGATGATACAATAGCAAGTTTCATTTCTTTTACCTTGAAACGTGAAAGTAA	3590
DB	2521	AAACAAACAGATGATACAATAGCAAGTTTCATTTCTTTTACCTTGAAACGTGAAAGTAA	2580
QY	3591	TCAGTGAGATCAGTTGCTTCTGCTTTTAAATTTTTCACAGTTTTCAGATTTTATGTTT	3650
DB	2581	TCAGTGAGATCAGTTGCTTCTGCTTTTAAATTTTTCACAGTTTTCAGATTTTATGTTT	2640
QY	3651	TTGGTTTCAATCACTTTTAAATATGCTGCTGCTCCCATCAAACTTGTTCAGACTTATAT	3710
DB	2641	TTGGTTTCAATCACTTTTAAATATGCTGCTGCTCCCATCAAACTTGTTCAGACTTATAT	2700
QY	3711	TTTCTAAAACTGGAGACAAACCAATAAATACAAAAAACAAGCTGCTTCTTCAAAATCT	3770
DB	2701	TTTCTAAAACTGGAGACAAACCAATAAATACAAAAAACAAGCTGCTTCTTCAAAATCT	2760
QY	3771	GCTGACTCAGATTCAGTGATCTTATAGCTTTTTCACAGCCAGACATAAATGGTGACG	3830
DB	2761	GCTGACTCAGATTCAGTGATCTTATAGCTTTTTCACAGCCAGACATAAATGGTGACG	2820
QY	3831	ATGATTAATGACAGTGACTATGATCCGCTGCTCCCATAGTGGGCTTTTAAATGTGATG	3890
DB	2821	ATGATTAATGACAGTGACTATGATCCGCTGCTCCCATAGTGGGCTTTTAAATGTGATG	2880
QY	3891	CATTGAAAAAGGACTATGTTTTCAGCTGTTTTCACAGTACTATGTTTATCTTTTA	3950
DB	2881	CATTGAAAAAGGACTATGTTTTCAGCTGTTTTCACAGTACTATGTTTATCTTTTA	2940
QY	3951	CCTATATTAGTGAATATCATTAGTAATCTTATCAATTTTAAATGTGACTGAAACC	4010
DB	2941	CCTATATTAGTGAATATCATTAGTAATCTTATCAATTTTAAATGTGACTGAAACC	3000
QY	4011	ATCCAGATCTGGAGTACCCCATCTTTCAAGAAATTAATGATAGTTTAAATTTGAG	4070
DB	3001	ATCCAGATCTGGAGTACCCCATCTTTCAAGAAATTAATGATAGTTTAAATTTGAG	3060
QY	4071	CTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTAAGTCAATGCACTTACTTTGCC	4130
DB	3061	CTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTAAGTCAATGCACTTACTTTGCC	3120
QY	4131	ATGGAATGAGAGAAATCAATAGATCAAGCTTATTAATCAATTTAACTTCAAGTCTTT	4190
DB	3121	ATGGAATGAGAGAAATCAATAGATCAAGCTTATTAATCAATTTAACTTCAAGTCTTT	3180

QY	4191	TTGCCATCTGCATATTTGGATTGGCAAGCTGTTGTTGATATCCCTTATTTTATCATTT	4250
DB	3181	TTGCCATCTGCATATTTGGATTGGCAAGCTGTTGTTGATATCCCTTATTTTATCATTT	3240
QY	4251	CTTATTTTGAATGCTAGGAAGCTTACTTGGCATTTTCATATTTGGATATATTTTATCTGTA	4310
DB	3241	CTTATTTTGAATGCTAGGAAGCTTACTTGGCATTTTCATATTTGGATATATTTTATCTGTA	3300
QY	4311	AAGTTCCTTCTGCTGCTGTTTTCCTTATTTGGTTATGTTTCCATCAGTATTTCTGTTCACT	4370
DB	3301	AAGTTCCTTCTGCTGCTGTTTTCCTTATTTGGTTATGTTTCCATCAGTATTTCTGTTCACT	3360
QY	4371	TATATTTGCTTCTTTCACCTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC	4430
DB	3361	TATATTTGCTTCTTTCACCTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC	3420
QY	4431	TATTTCTGTCGACGCTTGTGTTGATTTGCAATCACTGAAATAACTTTCTTTATGGGATAC	4490
DB	3421	TATTTCTGTCGACGCTTGTGTTGATTTGCAATCACTGAAATAACTTTCTTTATGGGATAC	3480
QY	4491	ACAATTTGCAACTTATTTCTTCAATATGCTTTTGTATCATCTTCAATCTATCCACTTTCTA	4550
DB	3481	ACAATTTGCAACTTATTTCTTCAATATGCTTTTGTATCATCTTCAATCTATCCACTTTCTA	3540
QY	4551	GGTTGCTGCTGATTTCTTTCATAAAGATTTCTTGAAGAAATGTCGAAAAATTTGGACACC	4610
DB	3541	GGTTGCTGCTGATTTCTTTCATAAAGATTTCTTGAAGAAATGTCGAAAAATTTGGACACC	3600
QY	4611	TATAATCCATTTGGGATAGGCTTTTTCAGTAGCTGTTTATATCGCTTACCTGCGATGTACTG	4670
DB	3601	TATAATCCATTTGGGATAGGCTTTTTCAGTAGCTGTTTATATCGCTTACCTGCGATGTACTG	3660
QY	4671	TGGATTTTCTCTTCTTACAATCTATGAGAAAAATATGGAGGAGATCAATAAGAAAGAT	4730
DB	3661	TGGATTTTCTCTTCTTACAATCTATGAGAAAAATATGGAGGAGATCAATAAGAAAGAT	3720
QY	4731	CCCTTTTTCAGAAACCTTTTCAACGAAAGTCTTAAATAATAGGAAGCTTCCAGAACCCAGAC	4790
DB	3721	CCCTTTTTCAGAAACCTTTTCAACGAAAGTCTTAAATAATAGGAAGCTTCCAGAACCCAGAC	3780
QY	4791	AATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGACTTAAAGGTCAAAGAGCTGATGGGT	4850
DB	3781	AATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGACTTAAAGGTCAAAGAGCTGATGGGT	3840
QY	4851	TGCGAGTGTGTGAGGAGAAACCATCCATTTATGCTCAGCAATTTTGCATTAAGAAATATGAT	4910
DB	3841	TGCGAGTGTGTGAGGAGAAACCATCCATTTATGCTCAGCAATTTTGCATTAAGAAATATGAT	3900
QY	4911	GACAAGAAAGATTTTCTTCTTTTCAAGAAAGTAAAGAAAGTGGCAACTTAAATACATCTCT	4970
DB	3901	GACAAGAAAGATTTTCTTCTTTTCAAGAAAGTAAAGAAAGTGGCAACTTAAATACATCTCT	3960
QY	4971	TTCTGTGTGAAAAAGGAGAGATCTTGGACTATTTGGGTCCTTAAATGGTGTGCGAAAGC	5030
DB	3961	TTCTGTGTGAAAAAGGAGAGATCTTGGGACTATTTGGGTCCTTAAATGGTGTGCGAAAGC	4020
QY	5031	ACAATTTTAAATATTTCTGCTGCTGATTTTGAACCAACTTTCAGGCCAGGATTTTATGGA	5090
DB	4021	ACAATTTTAAATATTTCTGCTGCTGATTTTGAACCAACTTTCAGGCCAGGATTTTATGGA	4080
QY	5091	GATTATTTCTTCAGAGCAAGTGAAGATGATTTTCACTGAAGTGTATGGGTACTGTCTCT	5150
DB	4081	GATTATTTCTTCAGAGCAAGTGAAGATGATTTTCACTGAAGTGTATGGGTACTGTCTCT	4140
QY	5151	CAGATAAAACCTTTTGTGGCCAGATCTACTATTCAGAGAAACATTTTGAAATTTATGAGCT	5210
DB	4141	CAGATAAAACCTTTTGTGGCCAGATCTACTATTCAGAGAAACATTTTGAAATTTATGAGCT	4200
QY	5211	GTCAAAGGAATGAGTCCAGTGCATGAAGAAAGTCAATAGTCCGAATAACACATGCACTT	5270
DB	4201	GTCAAAGGAATGAGTCCAGTGCATGAAGAAAGTCAATAGTCCGAATAACACATGCACTT	4260
QY	5271	GATTTTAAAGAAACATCTTTCAGAGAGACTGTAAAGAAACCTTACCTGCAGGAATCAAAACGAAAG	5330



Db 4261 GATTTAAAGAACATCTTCAGAGACTGTAAAGAACTCCTCAGGAATCAAAACGAAAG 4320  
Qy 5331 TTGTGTTTCTCTAAGTAGCTAGGGAATCCTCAGATTAATTTGCTAGATGAACCATCT 5390  
Db 4321 TTGTGTTTCTCTAAGTAGCTAGGGAATCCTCAGATTAATTTGCTAGATGAACCATCT 4380  
Qy 5391 ACAGGTATGATCCCAAGCCCAACAGCACATGTGGCGAGCAATTCGAATGCAATTTAA 5450  
Db 4381 ACAGGTATGATCCCAAGCCCAACAGCACATGTGGCGAGCAATTCGAATGCAATTTAA 4440  
Qy 5451 AACAGAAAGCGGCTGCTATTTCGACCACTCCTATATGAGGAGGAGCGGCTGCTGT 5510  
Db 4441 AACAGAAAGCGGCTGCTATTTCGACCACTCCTATATGAGGAGGAGCGGCTGCTGT 4500  
Qy 5511 GATCAGTAGCTATCAGTGTCTGGGAGGAGTGAAGTATCGAACAGTACAACTCTA 5570  
Db 4501 GATCAGTAGCTATCAGTGTCTGGGAGGAGTGAAGTATCGAACAGTACAACTCTA 4560  
Qy 5571 AAGAGTAAATTTGAAAGGCTACTTTTTCGAAATTAATTTGAAGGACTGATAGAAAC 5630  
Db 4561 AAGAGTAAATTTGAAAGGCTACTTTTTCGAAATTAATTTGAAGGACTGATAGAAAC 4620  
Qy 5631 CTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTCAG 5690  
Db 4621 CTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTCAG 4680  
Qy 5691 GAAAGTTTCTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTCAAGTCCCTTCA 5750  
Db 4681 GAAAGTTTCTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTCAAGTCCCTTCA 4740  
Qy 5751 CAATCTTTTATAGCTGGAAGAGCTAATACATGCTTTTCCCATTCGAAGATATAGCTTT 5810  
Db 4741 CAATCTTTTATAGCTGGAAGAGCTAATACATGCTTTTCCCATTCGAAGATATAGCTTT 4800  
Qy 5811 TCTCAAGCAATTTGGAACAGGTTTGTAGAACTCACTAAAGAACAAAGAGGAGGAAGAT 5870  
Db 4801 TCTCAAGCAATTTGGAACAGGTTTGTAGAACTCACTAAAGAACAAAGAGGAGGAAGAT 4860  
Qy 5871 AATAGTTGTGAATTTAAACAGACACATTTGTTGGGAACGACAAACAAAGATAGAGTA 5930  
Db 4861 AATAGTTGTGAATTTAAACAGACACATTTGTTGGGAACGACAAACAAAGATAGAGTA 4920  
Qy 5931 GTATTTTGA 5939  
Db 4921 GTATTTTGA 4929

RESULT 7

ABSS7750  
ID ABS57750 standard; cDNA; 4917 BP.  
XX AC ABS57750;  
XX AC ABS57750;  
XX DT 04-FEB-2003 (first entry)  
XX DE Coding sequence of human ATP binding cassette ABCA5 transporter #1.  
XX KW Human; ATP binding cassette; ABC; ABCA5; transporter;  
XX KW neurotoxin transport; beta-amyloid peptide; chromosome mapping;  
XX KW blood brain barrier transport; tissue typing; predictive medicine;  
XX KW ABCA5 mediated disorder; ABCA5 related disorder; gene therapy;  
XX KW gene; ss.  
XX OS Homo sapiens.  
XX FH Location/Qualifiers  
XX CDS 1..4917  
XX FT /\*tag= a  
XX FT /product= "ABCA5"  
XX FT /note= "ATP binding cassette (ABC) A5 transporter"  
XX FT /transl\_except= (pos:3160..3162, aa:Ser)  
XX FT

Query Match 73.7%; Score 4807.4; DB 25; Length 4917;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 4891; Conservative 0; Mismatches 7; Indels 50; Gaps 3;  
Qy 1011 ATGTCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACAGACACTTCTACTGAAG 1070  
Db 1 ATGTCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACAGACACTTCTACTGAAG 60  
Qy 1071 AATTACTTAATTAATGCAACCAACAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 1130  
Db 61 AATTACTTAATTAATGCAACCAACAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 120  
Qy 1131 TTTTCTTTTATTTGGTAAATTAATTAATAGCATGATGATCCCAATTAAGAAATATGAAGA 1190  
Db 121 TTTTCTTTTATTTGGTAAATTAATTAATAGCATGATGATCCCAATTAAGAAATATGAAGA 180  
Qy 1191 GTGCTTAATATAGAACTCAATCCTATGGCAAGTTTACTTCTTAATTAATTTCTTGA 1250  
Db 181 GTGCTTAATATAGAACTCAATCCTATGGCAAGTTTACTTCTTAATTTCTTGA 240  
Qy 1251 TATCTCCAGTGAATTAATTAATTAAGAGCATCATGAGAAAGTGCTACTGATCATCTA 1310  
Db 241 TATCTCCAGTGAATTAATTAATTAAGAGCATCATGAGAAAGTGCTACTGATCATCTA 300

QY	1311	CCTGATGTCATAATTA	1370
DB	301	CCTGATGTCATAATTA	360
QY	1371	TCTAAGCCGAGCAACT	1430
DB	361	TCTAAGCCGAGCAACT	420
QY	1431	TTTTTTTCCCTGATAT	1490
DB	421	TTTTTTTCCCTGATAT	480
QY	1491	AAATCATGTGAGGCTG	1550
DB	481	AAATCATGTGAGGCTG	540
QY	1551	GATGCTGCCATTAATA	1610
DB	541	GATGCTGCCATTAATA	600
QY	1611	AAAGCTGTTAATATGG	1670
DB	601	AAAGCTGTTAATATGG	660
QY	1671	ATTTTAAATATACCTA	1730
DB	661	ATTTTAAATATACCTA	720
QY	1731	GTAGCAGAAAAGAAAA	1790
DB	721	GTAGCAGAAAAGAAAA	780
QY	1791	GCCTTTTGGCTTTCCT	1850
DB	781	GCCTTTTGGCTTTCCT	840
QY	1851	ATGGCAGTCATGCGA	1910
DB	841	ATGGCAGTCATGCGA	900
QY	1911	CTGCTTTTTCCTTTAT	1970
DB	901	CTGCTTTTTCCTTTAT	960
QY	1971	TTTTAAAAATCAAACA	2030
DB	961	TTTTAAAAATCAAACA	1020
QY	2031	ATTGGCCTTATGATAA	2090
DB	1021	ATTGGCCTTATGATAA	1080
QY	2091	CCTTTCTGTCACTGTA	2150
DB	1081	CCTTTCTGTCACTGTA	1140
QY	2151	AATGAAGGTGCTTCAT	2210
DB	1141	AATGAAGGTGCTTCAT	1200
QY	2211	ATCATGCTCAACATTA	2270
DB	1201	ATCATGCTCAACATTA	1260
QY	2271	ATTCCAGGGAAATTTG	2330
DB	1261	ATTCCAGGGAAATTTG	1320
QY	2331	TCAAAGAGTAAAGAAAT	2390
DB	1321	TCAAAGAGTAAAGAAAT	1380

QY	2391	TTTAGTGAATAATTAAT	2450
DB	1381	TTTAGTGAATAATTAAT	1440
QY	2451	AGTGGTATTTCAGAAAC	2510
DB	1441	AGTGGTATTTCAGAAAC	1500
QY	2511	TCATTTTGACATATATC	2570
DB	1501	TCATTTTGACATATATC	1560
QY	2571	AGTACATTGATGAATA	2630
DB	1561	AGTACATTGATGAATA	1620
QY	2631	TATGGACACAGAGTCT	2690
DB	1621	TATGGACACAGAGTCT	1680
QY	2691	TGTCACACAGTATACAT	2750
DB	1681	TGTCACACAGTATACAT	1740
QY	2751	GCTTCAATCAAGGGGAT	2810
DB	1741	GCTTCAATCAAGGGGAT	1800
QY	2811	GATTTAGACATGACAGT	2870
DB	1801	GATTTAGACATGACAGT	1860
QY	2871	AGAAAGCTGTCAATTA	2930
DB	1861	AGAAAGCTGTCAATTA	1920
QY	2931	CCAAACAGCTGGAAAT	2990
DB	1921	CCAAACAGCTGGAAAT	1980
QY	2991	AAAGCCAAATCGGGTGA	3050
DB	1981	AAAGCCAAATCGGGTGA	2040
QY	3051	GATAGAAAGCTGTGAT	3110
DB	2041	GATAGAAAGCTGTGAT	2100
QY	3111	AAAAAGTAAATGGGGAT	3170
DB	2101	AAAAAGTAAATGGGGAT	2160
QY	3171	GAATCTCTTTCTTCACT	3230
DB	2161	GAATCTCTTTCTTCACT	2220
QY	3231	GACCAACAACTTGTGT	3290
DB	2221	GACCAACAACTTGTGT	2280
QY	3291	TCCTGCCCTAGACAGT	3350
DB	2281	TCCTGCCCTAGACAGT	2340
QY	3351	TTGGAAGACGTATTTT	3410
DB	2341	TTGGAAGACGTATTTT	2400
QY	3411	GTATTTTACTCAGCAGC	3470
DB	2401	GTATTTTACTCAGCAGC	2460
QY	3471	CAGAGCTTACTTATTT	3530

||||| 2461 CAGAGCTTACTTATCTTCTGGAACCAAGCGTCTAGTGAGACCAATGAGCGCTTTGG 2520  
Qy ||||| 3531 AACAACAGATGATACAAATAGCAAGATTTTCATTTCTTTACCTTGAAACGTGAAGTAAA 3590  
Db ||||| 2521 AACAACAGATGATACAAATAGCAAGATTTTCATTTCTTTACCTTGAAACGTGAAGTAAA 2580  
Qy ||||| 3591 TCAGTGAGATCAGTGTTGCTTCTGCTTTAAATTTTTCACAGTTCAGATTTTATGTTT 3650  
Db ||||| 2581 TCAGTGAGATCAGTGTTGCTTCTGCTTTAAATTTTTCACAGTTCAGATTTTATGTTT 2640  
Qy ||||| 3651 TTGGTTTCATCAGCTTTTAAATTTGCTGTTCCCATCAAACTGTTCCAGACTTATAT 3710  
Db ||||| 2641 TTGGTTTCATCAGCTTTTAAATTTGCTGTTCCCATCAAACTGTTCCAGACTTATAT 2700  
Qy ||||| 3711 TTCTTAAACCTGGAGCAAAACACATAAATACAAACAAAGTCTGCTTCTTCAAAATTC 3770  
Db ||||| 2701 TTCTTAAACCTGGAGCAAAACACATAAATACAAACAAAGTCTGCTTCTTCAAAATTC 2760  
Qy ||||| 3771 GCT-----GACTCAGATACAGTGATCTTATTAGCTTTTTCACAAGC 3812  
Db ||||| 2761 GCTGGTGAGATGNNGTGAAGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAAGC 2820  
Qy ||||| 3813 CAGAACATAATGCTGACGATGATTAATGACAGTGACTATGATCCGCTGCTCCCATAGT 3872  
Db ||||| 2821 CAGAACATAATGCTGACGATGATTAATGACAGTGACTATGATCCGCTGCTCCCATAGT 2880  
Qy ||||| 3873 GCGCTTTAAATGTGATGCAATCAGAAAAGGACTATGTTTTTTCGAGCTGTTTCAACAGT 3932  
Db ||||| 2881 GCGCTTTAAATGTGATGCAATCAGAAAAGGACTATGTTTTTTCGAGCTGTTTCAACAGT 2940  
Qy ||||| 3933 ACTATGCTTTTATCTTTTACCTATATTAGTGAATATCAATAGTAACTACTATCTTTATCAT 3992  
Db ||||| 2941 ACTATGCTTTTATCTTTTACCTATATTAGTGAATATCAATAGTAACTACTATCTTTATCAT 3000  
Qy ||||| 3993 TTAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTTACTGAT 4052  
Db ||||| 3001 TTAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTTACTGAT 3060  
Qy ||||| 4053 ATAGTTTTTAAATGTAGCTGATTTTCAAGCAGCTTTGCTGG- AATCATTTGTTACTGC 4111  
Db ||||| 3061 ATAGTTTTTAAATGTAGCTGATTTTCAAGCAGCTTTGCTGGAAATCATTTGAACTGC 3120  
Qy ||||| 4112 AATGCCACCTTACTTTGCCATCTGCATATTTGGAATGGACAAGCTGTTTGTGATAT 4171  
Db ||||| 3121 AATGCCACCTTACTTTGCCATCTGCATATTTGGAATGGACAAGCTGTTTGTGATAT 3161  
Qy ||||| 4172 ACTTAAACTTTTTCAGGCTCTTTTGGCATTTGGAATGGACAAGCTGTTTGTGATAT 4231  
Db ||||| 3162 -----AGGCTCTTTTGGCATCTGCATATTTGGAATGGACAAGCTGTTTGTGATAT 3209  
Qy ||||| 4232 CCCCTTATTTTATCAATCTTATTTTGTGCTGAGGAGCTTACTGGCAATTCATTTATGG 4291  
Db ||||| 3210 CCCCTTATTTTATCAATCTTATTTTGTGCTGAGGAGCTTACTGGCAATTTCAITATGG 3269  
Qy ||||| 4292 ATTATATTTTATCTGAAAGTTCCTGCTGCTGTTTTTGGCTTATTTGTTTATGTTCC 4351  
Db ||||| 3270 ATTATATTTTATCTGAAAGTTCCTGCTGCTGTTTTTGGCTTATTTGCTTATGTTTCC 3329  
Qy ||||| 4352 ATCAGTTTATTTCTGTTCTTATTTGCTTCTTTTCAACCTTTTAAAGAAATTTTAAATACCAA 4411  
Db ||||| 3330 ATCAGTTTATTTCTGTTCTTATTTGCTTCTTTTCAACCTTTTAAAGAAATTTTAAATACCAA 3389  
Qy ||||| 4412 AGAATTTTGTGATTTATCTTCTGCTGAGGAGCTTGNCTTGTATTTGCAATCTCAAT 4471  
Db ||||| 3390 AGAATTTTGTGATTTATCTTCTGCTGAGGAGCTTGNCTTGTATTTGCAATCTCAAT 3449  
Qy ||||| 4472 AACTTTCTTTATGGATACAAATTTGCAATTTCTTCAATTTGCTTTGCTTTGATCATCAT 4531  
Db ||||| 3450 AACTTTCTTTATGGATACAAATTTGCAATTTCTTCAATTTGCTTTTGTGATCATCAT 3509  
Qy ||||| 4532 TCCAATCTATCCACTTCTAGGTTGCTGATTTCTTTTCAATAAGATTTCTTTGGAAGATGT 4591  
Db |||||

Db ||||| 3510 TCCAATCTATCCACTTCTAGGTTGCTGATTTCTTTTCAATAAGATTTTCTTGGAAAGATGT 3569  
Qy ||||| 4592 ACGAAAAATGTGGACACCTATAATCAATGGATAGGCTTTTCAGTAGCTGTTTATATGCC 4651  
Db ||||| 3570 ACGAAAAATGTGGACACCTATAATCAATGGATAGGCTTTTCAGTAGCTGTTTATATGCC 3629  
Qy ||||| 4652 TTACCTGAGTGTGTGCTGCTGCTTCTTCAATATCTATGAGAAAAAATATGGAGG 4711  
Db ||||| 3630 TTACCTGAGTGTGTGCTGCTGCTTCTTCAATATCTATGAGAAAAAATATGGAGG 3689  
Qy ||||| 4712 CAGATCAATATAAGAAAGATCCCTTTTTCAGAAAACTTTTCAACGAAGTCTTAAATATAGAA 4771  
Db ||||| 3690 CAGATCAATATAAGAAAGATCCCTTTTTCAGAAAACTTTTCAACGAAGTCTTAAATATAGAA 3749  
Qy ||||| 4772 GCTTCCAGAACACACAGACAATGAGGATGAAGATGAAGATGTCAAGCTGAAAGCTAAA 4831  
Db ||||| 3750 GCTTCCAGAACACACAGACAATGAGGATGAAGATGAAGATGTCAAGCTGAAAGCTAAA 3809  
Qy ||||| 4832 GGTCAAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAACCATCCATTTATGGTCAGCAA 4891  
Db ||||| 3810 GGTCAAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAACCATCCATTTATGGTCAGCAA 3869  
Qy ||||| 4892 TTTGCAATAAGAAATATGATGACAGAAAGATTTTCTTCTTCAAGAAAAAGTAAAGAAAT 4951  
Db ||||| 3870 TTTGCAATAAGAAATATGATGACAGAAAGATTTTCTTCTTCAAGAAAAAGTAAAGAAAT 3929  
Qy ||||| 4952 GCGAACTAAATACATCTCTTCTGTGTAAGAAAGGAGAGATCTTGGAGCTATTTGGGTCC 5011  
Db ||||| 3930 GCGAACTAAATACATCTCTTCTGTGTAAGAAAGGAGAGATCTTGGAGCTATTTGGGTCC 3989  
Qy ||||| 5012 AAATGGTCTGGCAAAAGCACAATTTATTAATTTCTGTTGGTGATATTTGAACCAACTTC 5071  
Db ||||| 3990 AAATGGTCTGGCAAAAGCACAATTTATTAATTTCTGTTGGTGATATTTGAACCAACTTC 4049  
Qy ||||| 5072 AGGCCAGGATTTTATGAGAGATTTCTTTCAGAGCAAGTGAAGATGATGATCTACTGAA 5131  
Db ||||| 4050 AGGCCAGGATTTTATGAGAGATTTCTTTCAGAGCAAGTGAAGATGATGATCTACTGAA 4109  
Qy ||||| 5132 GTGTATGGGTTACTGCTCTCAGATAAACCCCTTTGTCGAGATCTACTACATTTGCAGGAACA 5191  
Db ||||| 4110 GTGTATGGGTTACTGCTCTCAGATAAACCCCTTTGTCGAGATCTACTACATTTGCAGGAACA 4169  
Qy ||||| 5192 TTTTGAATTTTATGAGAGCTGTCAAAGGAATGAGTGCAGATGACATGAAAGAGTCTAAG 5251  
Db ||||| 4170 TTTTGAATTTTATGAGAGCTGTCAAAGGAATGAGTGCAGATGAAAGAGTCTAAG 4229  
Qy ||||| 5252 TCGAATAACACATGCACTTTGATTTAAAGAACATCTTTCAGAAAGCTGTAAAGAACTACC 5311  
Db ||||| 4230 TCGAATAACACATGCACTTTGATTTAAAGAACATCTTTCAGAAAGCTGTAAAGAACTACC 4289  
Qy ||||| 5312 TCGAGGAATCAAAAGAAAGTTGTTTGTCTTAAAGTATGCTAGGGAATCTTCAAGATTAC 5371  
Db ||||| 4290 TCGAGGAATCAAAAGAAAGTTGTTTGTCTTAAAGTATGCTAGGGAATCTTCAAGATTAC 4349  
Qy ||||| 5372 TTTTGTAGATGAACCATCTACAGGATATGATCCCAAAGCCAAACAGCAATGTGCGGAGC 5431  
Db ||||| 4350 TTTTGTAGATGAACCATCTACAGGATATGATCCCAAAGCCAAACAGCAATGTGCGGAGC 4409  
Qy ||||| 5432 AATTCGAACCTGATTTTAAAGAACAGAAAGGGGCTCTATTTCTGACCACTCATATATGGA 5491  
Db ||||| 4410 AATTCGAACCTGATTTTAAAGAACAGAAAGGGGCTCTATTTCTGACCACTCATATATGGA 4469  
Qy ||||| 5492 GGAGGACAGAGGCTGCTGTGATCTGAGTAGCTATCATGTTGCTGCGGAGTAAAGATGAT 5551  
Db ||||| 4470 GGAGGACAGAGGCTGCTGTGATCTGAGTAGCTATCATGTTGCTGCGGAGTAAAGATGAT 4529  
Qy ||||| 5552 CGGAACAGTACAACTATAAGAGTAAATTTTGGAAAAAGGCTACTTTTGGAAATTTAAAT 5611  
Db ||||| 4530 CGGAACAGTACAACTATAAGAGTAAATTTTGGAAAAAGGCTACTTTTGGAAATTTAAAT 4589  
Qy ||||| 5612 GAAGGACCTGGATAGAAAAACCTAGAAAGTAGACCGCCCTTCAAGAGAAATTCAGTATATTT 5671  
Db ||||| 4590 GAAGGACCTGGATAGAAAAACCTAGAAAGTAGACCGCCCTTCAAGAGAAATTCAGTATATTT 4649

```
QY 5672 CCAGATGCAAGCGTCAGGAAAGTTTCTTCTATTTTGGCTTATATAAATTCCTAAGGA 5731
DB 4650 CCAGATGCAAGCGTCAGGAAAGTTTCTTCTATTTTGGCTTATATAAATTCCTAAGGA 4709
QY 5732 AGATGTCAGTCCTCTTCAACAATCTTTTAAAGCTGGAAGAGCTTAAACATGCTTTTGC 5791
DB 4710 AGATGTCAGTCCTCTTCAACAATCTTTTAAAGCTGGAAGAGCTTAAACATGCTTTTGC 4769
QY 5792 CATTGAAGATATAGCTTTTCTCAAGCAACATTGGACACAGTCTTTTGTAGAACTCACTAA 5851
DB 4770 CATTGAAGATATAGCTTTTCTCAAGCAACATTGGACACAGTCTTTTGTAGAACTCACTAA 4829
QY 5852 AGAACAGAGGAGGAAGATATAGTTGTGGAACCTTTTAAACAGCACACACTTTTGGTGGGAACG 5911
DB 4830 AGAACAGAGGAGGAAGATATAGTTGTGGAACCTTTTAAACAGCACACACTTTTGGTGGGAACG 4889
QY 5912 AACACAGAGATAGAGTAGTATTTTGA 5939
DB 4890 AACACAGAGATAGAGTAGTATTTTGA 4917

RESULT 8
AAD37619
ID AAD37619 standard; cDNA; 4785 BP.
XX
AC AAD37619;
XX
DT 10-SEP-2002 (first entry)
XX
DE Human transporter protein cDNA #2.
XX
KW Human; novel human protein; NHP; transporter protein; mental disorder;
KW cancer; gene therapy; drug screening; nutraceutical application; gene;
KW cosmetic application; polymorphism; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..4785
FT /*tag= a
FT /product= "Human transporter protein #2"
FT /transl_except= (pos:808..810, aa:Tyr)
FT /transl_except= (pos:2494..2496, aa:Ser)
FT /transl_except= (pos:2878..2880, aa:Val)
FT misc_feature 810
FT /*tag= b
FT /note= "This degenerate base represents a polymorphic site"
FT misc_feature 2494
FT /*tag= c
FT /note= "This degenerate base represents a polymorphic site"
FT misc_feature 2878
FT /*tag= d
FT /note= "This degenerate base represents a polymorphic site"
XX
PN WO200231147-A2.
XX
PD 18-APR-2002.
XX
XX 04-OCT-2001; 2001WO-US31113.
XX
PR 10-OCT-2000; 2000US-239629P.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Hu Y, Nepomnichy B;
XX
XX WPI; 2002-452347/48.
XX
XX P-PSDB; AAE23657.
XX
```

```
PT Novel nucleic acid molecule encoding novel human proteins, useful for
XX therapeutic, diagnostic and pharmacogenomic applications
PS Claim 1; Page 39-41; 46pp; English.
XX
CC The present sequence is a cDNA coding for novel human protein (NHP),
CC human transporter protein. NHPs shares structural similarity with
CC the mammalian ATP-binding cassette (ABC) transporters and multidrug
CC resistance transporters. NHP polynucleotides are useful for the
CC therapeutic, diagnostic and pharmacogenomic applications. They are
CC used for detecting and treating mental disorders and cancers. They
CC are also used in gene therapy. NHP polypeptides are useful for
CC diagnosis, drug screening, clinical trial monitoring, treatment of
CC diseases and disorders, and cosmetic or nutraceutical applications.
XX
SQ Sequence 4785 BP; 1489 A; 807 C; 917 G; 1568 T; 4 other;
```

```
Query Match 72.9%; Score 4754.8; DB 24; Length 4785;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 4756; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1011 ATGTCCTCACTGCAATTAGGGAGGTAGGAGTTTGGAGACAGACAGACACTTCTACTGAAG 1070
DB 1 ATGTCCTCACTGCAATTAGGGAGGTAGGAGTTTGGAGACAGACAGACACTTCTACTGAAG 60
QY 1071 AATTACTTAAATTAATGCGAAGCAAAAAAGAGTAGTGTTCAGGAAATTTCTTTTCCACTA 1130
DB 61 AATTACTTAAATTAATGCGAAGCAAAAAAGAGTAGTGTTCAGGAAATTTCTTTTCCACTA 120
QY 1131 TTTTCTTTTATTTGGTTAAATTAATTAATGATGATGATCCAAATAAGAAATATGAAGAA 1190
DB 121 TTTTCTTTTATTTGGTTAAATTAATTAATGATGATGATCCAAATAAGAAATATGAAGAA 180
QY 1191 GTGCCTTAATATAGAACTCAATCTATGGAACAGTTTACTCTTTCTTAATCTTATCTTGA 1250
DB 181 GTGCCTTAATATAGAACTCAATCTATGGAACAGTTTACTCTTTCTTAATCTTATCTTGA 240
QY 1251 TATACTCCAGTGACTAATTAATTAACAGCAGCATCATGAGAAAGTGTCTACTGATCATCTA 1310
DB 241 TATACTCCAGTGACTAATTAATTAACAGCAGCATCATGAGAAAGTGTCTACTGATCATCTA 300
QY 1311 CTTGATGTCATTAATTAATGAGAAATATACAAATGAAAGAAATGTTAAACATCAAGTCTC 1370
DB 301 CTTGATGTCATTAATTAATGAGAAATATACAAATGAAAGAAATGTTAAACATCAAGTCTC 360
QY 1371 TCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAAAGACTCCCATGCTTATGACTTCGT 1430
DB 361 TCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAAAGACTCCCATGCTTATGACTTCGT 420
QY 1431 TTTTCTTCTGATATGATTCAGTATCTTCTATTTATATGATTCAGAGCTGGCTGTTC 1490
DB 421 TTTTCTTCTGATATGATTCAGTATCTTCTATTTATATGATTCAGAGCTGGCTGTTC 480
QY 1491 AAATCAATGTCAGGCTGCTCAGTACTCGTCTCAGGTTTCAAGTTTCAAGCATCCATA 1550
DB 481 AAATCAATGTCAGGCTGCTCAGTACTCGTCTCAGGTTTCAAGTTTCAAGCATCCATA 540
QY 1551 GATGCTGCCATTTATACAGTTGAAGCAATGTTTCTTTTGGAGAGCTGGAGTCAACT 1610
DB 541 GATGCTGCCATTTATACAGTTGAAGCAATGTTTCTTTTGGAGAGCTGGAGTCAACT 600
QY 1611 AAAGCTGTTTATATGAGGAGAACTGCTGTGTGTAAGAAATAGATACCTTTTCCCGAGGAGTA 1670
DB 601 AAAGCTGTTTATATGAGGAGAACTGCTGTGTGTAAGAAATAGATACCTTTTCCCGAGGAGTA 660
QY 1671 ATTTTAATATACCTAGTATAGCAATTTTCACTTTTGGATATCTTTTGGCAATTCATATC 1730
DB 661 ATTTTAATATACCTAGTATAGCAATTTTCACTTTTGGATATCTTTTGGCAATTCATATC 720
QY 1731 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTTCATGATATC 1790
DB 721 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTTCATGATATC 780
```



Db 2941 CCTATATTAGTAATATCATATTAGTAACATACTAATCTTTATCATTTAAATGAGTGAAC 3000  
Qy 4011 ATCCAGATCTGGAGTACCCCAATCTTTCAAGAAATTAAGTATAGTATTTTAAATTTGAG 4070  
Db 3001 ATCCAGATCTGGAGTACCCCAATCTTTCAAGAAATTAAGTATAGTATTTTAAATTTGAG 3060  
Qy 4071 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCCACCTTACTTTGCC 4130  
Db 3061 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCCACCTTACTTTGCC 3120  
Qy 4131 ATGMAAATCGAGAGATCATAGATCAAGAGCTTATCTCAACTTAACTTTTCAAGTCTTT 4190  
Db 3121 ATGMAAATCGAGAGATCATAGATCAAGAGCTTATCTCAACTTAACTTTTCAAGTCTTT 3180  
Qy 4191 TTGCAATCTGCATATTGGATTGGACAAGCTTTGTTGATATCCCTTATTTTATCATTT 4250  
Db 3181 TTGCAATCTGCATATTGGATTGGACAAGCTTTGTTGATATCCCTTATTTTATCATTT 3240  
Qy 4251 CTTATTTTGAATGCTAGGAAGCTTACTGCAATTTCAATTTATGGAATATATTTTATPACTGTA 4310  
Db 3241 CTTATTTTGAATGCTAGGAAGCTTATTTGGCAATTTCAATTTATGGAATATATTTTATPACTGTA 3300  
Qy 4311 AAGTTCCTCTGCTGGTCTTTTGGCTTATTTGGTATGTTCCATCAGTATTTCTGTTCACT 4370  
Db 3301 AAGTTCCTCTGCTGGTCTTTTGGCTTATTTGGTATGTTCCATCAGTATTTCTGTTCACT 3360  
Qy 4371 TATATTGCTTCTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 4430  
Db 3361 TATATTGCTTCTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3420  
Qy 4431 TATTCTGTGCGAGCGTTGCTTGTATTGCAATCACTGAAATAAATTTCTTTTATGGGATAC 4490  
Db 3421 TATTCTGTGCGAGCGTTGCTTGTATTGCAATCACTGAAATAAATTTCTTTTATGGGATAC 3480  
Qy 4491 ACAATTCGAATATTCTTCATATTATGCTTTTGTATCATCATTTCCAATCTATCCACTCTA 4550  
Db 3481 ACAATTCGAATATTCTTCATATTATGCTTTTGTATCATCATTTCCAATCTATCCACTCTA 3540  
Qy 4551 GGTGCTCTGATTTCTTTTCAATAAGATTTCTTTGGAAGAAATGTACGAAAAAATTTGGACACC 4610  
Db 3541 GGTGCTCTGATTTCTTTTCAATAAGATTTCTTTGGAAGAAATGTACGAAAAAATTTGGACACC 3600  
Qy 4611 TATAATCCATGGGATAGGCTTTTCAGTAGCTTTTATATCGCTTTACCTGAGTGTGACTG 4670  
Db 3601 TATAATCCATGGGATAGGCTTTTCAGTAGCTTTTATATCGCTTTACCTGAGTGTGACTG 3660  
Qy 4671 TGGATTTTCTCTTTACAATACTATGAGAAAAAATATGAGAGCAGATCAATAAGAAAAAGAT 4730  
Db 3661 TGGATTTTCTCTTTACAATACTATGAGAAAAAATATGAGAGCAGATCAATAAGAAAAAGAT 3720  
Qy 4731 CCTTTTTCAGAAACCTTTTCAACGAAGCTTAAAAATAGGAAGCTTCCAGAACCCACAGAC 4790  
Db 3721 CCTTTTTCAGAAACCTTTTCAACGAAGCTTAAAAATAGGAAGCTTCCAGAACCCACAGAC 3780  
Qy 4791 AATGAGGATGAAGATGAAGATGTCAAAAGCTGAAAGACTAAAGGCTCAAGAGCTGATGGGT 4850  
Db 3781 AATGAGGATGAAGATGAAGATGTCAAAAGCTGAAAGACTAAAGGCTCAAGAGCTGATGGGT 3840  
Qy 4851 TGGCAGTGTGTGAGGAGAAACCATTCATTTATGCTCAGCAATTTTGCATAAAGAAATATGAT 4910  
Db 3841 TGGCAGTGTGTGAGGAGAAACCATTCATTTATGCTCAGCAATTTTGCATAAAGAAATATGAT 3900  
Qy 4911 GACAAGAAAGATTTCTTTCTTTCAAGAAAGTGAAGAAAGTGGCACTTAATACATCTCT 4970  
Db 3901 GACAAGAAAGATTTCTTTCTTTCAAGAAAGTGAAGAAAGTGGCACTTAATACATCTCT 3960  
Qy 4971 TTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGGTCCAATGTGCTGGCAAAAGC 5030  
Db 3961 TTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGGTCCAATGTGCTGGCAAAAGC 4020  
Qy 5031 ACAATTTAATATTCTGGTTGGTGATATTGAACCAACTTCAGGCCAGGTATTTTATAGGA 5090  
Db 4021 ACAATTTAATATTCTGGTTGGTGATATTGAACCAACTTCAGGCCAGGTATTTTATAGGA 4080

Qy 5091 GATTATTTCTTTCAGAGACAAGTGAAGATGATGATTTCACTCAAGTGTATGGGTTACTGTCTCT 5150  
Db 4081 GATTATTTCTTTCAGAGACAAGTGAAGATGATGATTTCACTCAAGTGTATGGGTTACTGTCTCT 4140  
Qy 5151 CAGATAAAACCTTTTGTGGCCAGATACATCATTCGAGAAACATTTTGAATTTATGGAGCT 5210  
Db 4141 CAGATAAAACCTTTTGTGGCCAGATACATCATTCGAGAAACATTTTGAATTTATGGAGCT 4200  
Qy 5211 GTCAAAAGGAATGAGTCAAAAGTGCATGAAAGAGTGCATAGTGCATATACACATGCACTT 5270  
Db 4201 GTCAAAAGGAATGAGTCAAAAGTGCATGAAAGAGTGCATAGTGCATATACACATGCACTT 4260  
Qy 5271 GATTTAAAAGAAACATCTTTCAGAGAGCTGTAAAGAAACCTACCTCGAGAAATCAAAACGAAAG 5330  
Db 4261 GATTTAAAAGAAACATCTTTCAGAGAGCTGTAAAGAAACCTACCTCGAGAAATCAAAACGAAAG 4320  
Qy 5331 TTGTGTTTGTCTTAAGTATGCTAGGGAATCTTCAGATTAATCTTTGCTAGATGAACCATCT 5390  
Db 4321 TTGTGTTTGTCTTAAGTATGCTAGGGAATCTTCAGATTAATCTTTGCTAGATGAACCATCT 4380  
Qy 5391 ACAGGTATGATGCCCAAGCCAAAGCAGACATGTCGAGCAATTCGAACTGCAATTTAAA 5450  
Db 4381 ACAGGTATGATGCCCAAGCCAAAGCAGACATGTCGAGCAATTCGAACTGCAATTTAAA 4440  
Qy 5451 AACAGAAAGCGGCTGCTATTCTGACCACTCCTATATGAGGAGGAGGAGGCTGTCTGT 5510  
Db 4441 AACAGAAAGCGGCTGCTATTCTGACCACTCCTATATGAGGAGGAGGAGGCTGTCTGT 4500  
Qy 5511 GATCGAGTAGCTATCATGTTGCTGCGCAGTTAAGATGTATCGAAACAGTACAACTCTA 5570  
Db 4501 GATCGAGTAGCTATCATGTTGCTGCGCAGTTAAGATGTATCGAAACAGTACAACTCTA 4560  
Qy 5571 AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTTGAAGAGCTGATAGAAAC 5630  
Db 4561 AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTTGAAGAGCTGATAGAAAC 4620  
Qy 5631 CTAGAAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTCAG 5690  
Db 4621 CTAGAAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTCAG 4680  
Qy 5691 GAAAGTTTTCCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTCAAGTCCCTTTCA 5750  
Db 4681 GAAAGTTTTCCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTCAAGTCCCTTTCA 4740  
Qy 5751 CAATCTTTTAAAGCTGGAAGAGCT 5777  
Db 4741 CAATCTTTTAAAGCTGGAAGAGCT 4767

## RESULT 9

AB235938

ID AB235938 standard; cDNA; 2481 BP.

XX AC AB235938;

XX AC AC

DT 10-FEB-2003 (first entry)

XX Human secretory polynucleotide SPTM SEQ ID NO 102.

Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;  
asthma; Crohn's disease; neurological disorder; epilepsy; cancer;  
Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;  
multiple sclerosis; Parkinson's disease; cell proliferative disorder;  
anti-inflammatory; immunosuppressive; neuroprotective; nontropic;  
neuroleptic; anticonvulsant; cytosolic; antiparkinsonian; anxiolytic;  
antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;  
secretory polynucleotide; secretory protein; gene; ss.

OS Homo sapiens.

XX XX

PN WO200283876-A2.

XX XX





```
Db 1407 ATAAATTACTGAAGATATACAAATGAAAAAGAAATGTAATCAATCCAGTCTCTCTAAAGCCG 1466
QY 1380 AGCAACTTTCTAGTGTGGTTTTCAAGAGCTCCATGTCCTATGAACCTTCGTTTTTTTCTT 1439
Db 1467 AGCACTTTGTAGTGTGGTTTTCAAGAGCTCCATGTCCTATGAACCTTCGTTTTTTTCTT 1526
QY 1440 GATATGATCCAGTATCTTCTATTATATGATGATCAAGAGCTGGCTGTTCAAAATCATGT 1499
Db 1527 GATATGATCCAGTATCTTCTATTATATGATGATCAAGAGCTGGCTGTTCAAAATCATGT 1586
QY 1500 GAGGCTGCTCAGTACTGCTCTCAGTTTCAAGTTTCAAGCATCCATAGATGCTGCC 1559
Db 1587 GAGGCTGCTCAGTACTGCTCTCAGTTTCAAGTTTCAAGCATCCATAGATGCTGCC 1646
QY 1560 ATTATACAGTTGAAGACCAATGTTCTCTTTGGAAGGAGCTGGAGTCAACGTAAGCTGTT 1619
Db 1647 ATTATACAGTTGAAGACCAATGTTCTCTTTGGAAGGAGCTGGAGTCAACGTAAGCTGTT 1706
QY 1620 ATTATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTTCCCGAGGAGTAATTTTAATA 1679
Db 1707 ATTATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTTCCCGAGGAGTAATTTTAATA 1766
QY 1680 TACCTAGTTATAGCATTTTCACTTTTGGATACCTTTTGGCAATTCATATCGTAGCAGAA 1739
Db 1767 TACCTAGTTATAGCATTTTCACTTTTGGATACCTTTTGGCAATTCATATCGTAGCAGAA 1826
QY 1740 AAAGAAAAAAATAAAGAAATTTTAAAGATATGGGAGCTTCATGATACCTGCTTTTGG 1799
Db 1827 AAAGAAAAAAATAAAGAAATTTTAAAGATATGGGAGCTTCATGATACCTGCTTTTGG 1886
QY 1800 CTTTCTCGGTTCTTCTATATACAAAGTTTAAATTTTCTTATGTCCTTTCTTATGGCAGTC 1859
Db 1887 CTTTCTCGGTTCTTCTATATACAAAGTTTAAATTTTCTTATGTCCTTTCTTATGGCAGTC 1946
QY 1860 ATTGCAGACGCTCTTTGTTATTTTCTCAAAGTAGCAGCATTTGTGATATTTCTGTTTTT 1919
Db 1947 ATTGCAGACGCTCTTTGTTATTTTCTCAAAGTAGCAGCATTTGTGATATTTCTGTTTTT 2006
QY 1920 TTCTTTATGATATCATCTGATTTTGTGCTTTTAAATGCTGACACTCTTTTAAATAA 1979
Db 2007 TTCTTTATGATATCATCTGATTTTGTGCTTTTAAATGCTGACACTCTTTTAAATAA 2066
QY 1980 TCAAAACATGTGGAAATAGTTGAATTTTGTACTGTGCTTTTGGATTTATTTGGCCTT 2039
Db 2067 TCAAAACATGTGGAAATAGTTGAATTTTGTACTGTGCTTTTGGATTTATTTGGCCTT 2126
QY 2040 ATGATTAATCTCATAGAAATTTTCCAAATCGTTAGTGTGGCTTTTCAGTCTCTTTCTGT 2099
Db 2127 ATGATTAATCTCATAGAAATTTTCCAAATCGTTAGTGTGGCTTTTCAGTCTCTTTCTGT 2186
QY 2100 CACTGTACTTTTCTGATTTGTTATGTCACAGGTCTATGCTTATGATCAAGTCAATTTAAGGT 2159
Db 2187 CACTGTACTTTTCTGATTTGTTATGTCACAGGTCTATGCTTATGATCAAGTCAATTTAAGGT 2246
QY 2160 GCTTCAATTTTCAATTTTGAAGTGTGAGGAGGATATCTTAAATTTTAAATTTATCATGCTC 2219
Db 2247 GCTTCAATTTTCAATTTTGAAGTGTGAGGAGGATATCTTAAATTTTAAATTTATCATGCTC 2306
QY 2220 ACATTTAATAGTATATTTCTATGTCCTTTGGCTGTCTATCTTATGATCAAGTCAATTTCAAGG 2279
Db 2307 ACATTTAATAGTATATTTCTATGTCCTTTGGCTGTCTATCTTATGATCAAGTCAATTTCAAGG 2366
QY 2280 GAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAGGCTTCATATGGTCAAGAGT 2339
Db 2367 GAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAGGCTTCATATGGTCAAGAGT 2426
QY 2340 AAAAGAAATTTATCAGGAGTTTATCAGAGGGCAATGTTAAATGGAATATTTAGTTTTTA 2394
Db 2427 AAAAGAAATTTATCAGGAGTTTATCAGAGGGCAATGTTAAATGGAATATTTAGTTTTTA 2481
```

```
ABK35707
ID ABK35707 standard; cDNA; 3928 BP.
XX
AC ABK35707;
XX
DT 08-MAY-2002 (first entry)
XX
cDNA sequence #98 encoding novel human secreted protein.
DE
DE Human secreted protein; hyperproliferative disorder; autoimmune disorder;
XX immune deficiency disorder; blood disorder; inflammatory disorder;
XX infectious disorder; allergic condition; neurodegenerative disorder;
KW liver fibrosis; coagulation disorder; gene therapy; antimicrobial;
KW tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200177289-A2.
XX
XX 18-OCT-2001.
XX
XX 29-MAR-2001; 2001WO-US10232.
XX
XX 06-APR-2000; 2000US-195605P.
XX
XX (GEMY ) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GG;
PI Clark HF, Fectel K, Howes SH, Resnick RJ, Gulukota K, Graham JR;
XX
XX WPI; 2002-179322/23.
XX
XX Six hundred and twenty three polynucleotides derived from a variety of
PT human tissue sources which encode secreted proteins, useful for
PT treating immune deficiencies and disorders such as autoimmune disorders
XX
XX Claim 1; Page 127-128; 393pp; English.
XX
XX The present invention relates to the isolation of novel cDNA sequences
CC which encode human secreted proteins. The cDNA sequences have been
CC derived from a variety of human tissues. The invention also provides
CC a method for producing proteins from these polynucleotide sequences.
CC The proteins are useful for identifying compounds that modulate their
CC activity and production. The sequences of the invention are
CC useful for treating diseases such as hyperproliferative disorders
CC (e.g. cancer), immune deficiency disorders (e.g. severe combined
CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
CC sclerosis), blood disorders (e.g. thrombocytopaenia), inflammatory
CC disorders (e.g. arthritis), infectious disorders (e.g. hepatitis),
CC allergic conditions (e.g. asthma), neurodegenerative disorders (e.g.
CC Alzheimer's disease), liver fibrosis, coagulation disorders (e.g.
CC haemophilia), and tumours. The polynucleotide sequences of the
CC invention are also useful in gene therapy. ABK35610-ABK36232 represent
CC the cDNA sequences of the invention that encode for novel human
CC secreted proteins.
XX
XX Sequence 3928 BP; 1139 A; 681 C; 799 G; 1309 T; 0 other;
```

```
Query Match 27.4%; Score 1789.4; DB 24; Length 3928;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1833; Conservative 0; Mismatches 1; Indels 33; Gaps 1;
```

```
QY 1941 GATATTTTGTCTTTAATGTCACACTCTTTTAAAAAATCAAAACATGTGGGAATAGTT 2000
Db 1 GATATTTTGTCTTTAATGTCACACTCTTTTAAAAAATCAAAACATGTGGGAATAGTT 60
QY 2001 GAATTTTGTCTTACTGTGGCTTTTGGATTTATTTGGCTTATGATATCTCATAGAAGT 2060
Db 61 GAATTTTGTCTTACTGTGGCTTTTGGATTTATTTGGCTTATGATATCTCATAGAAGT 120
QY 2061 TTTCCCAAAATCGTTAGTGTGGCTTTTTCAGTCTCTTCTGTCACTGTACTTTTGTGTTGGT 2120
```

121 TTTCCCAATCGTAGTGGCTTTTCAGTCCCTTCTGTCACGTACTTTTGTGATTGGT 180  
2121 ATTGCACAGGTCAATGATTTAGAGATTTTAAAGAGGTGCTTCATTTTCAAAATTTGACT 2180  
181 ATTGCACAGGTCAATGATTTAGAGATTTTAAAGAGGTGCTTCATTTTCAAAATTTGACT 240  
2181 GCAGGCCCATATCTCTAAATTTACAAATTTATCATCTCATCTCATCTATATATCTAT 2240  
241 GCAGGCCCATATCTCTAAATTTACAAATTTATCATCTCATCTCATCTATATATCTAT 300  
2241 GTCTCTTGTGCTGTCTATCTTATCAAGTCAATTTCCAGGGGAATTTGGCTTTACGGAGATCA 2300  
301 GTCTCTTGTGCTGTCTATCTTATCAAGTCAATTTCCAGGGGAATTTGGCTTTACGGAGATCA 360  
2301 TCTTTATATTTTCTGAAGCCTTCATATTTGTCTCAAGAGTAAAGAAATTTATGAGGAGTTA 2360  
361 TCTTTATATTTTCTGAAGCCTTCATATTTGTCTCAAGAGTAAAGAAATTTATGAGGAGTTA 420  
2361 TCAGAGGCCAATGTTAATGGAATATTTAGTTTGTGAAATTTATGAGCCAGTTCTTCTCA 2420  
421 TCAGAGGCCAATGTTAATGGAATATTTAGTTTGTGAAATTTATGAGCCAGTTCTTCTCA 480  
2421 GAAATTTGTAGGAAAGCAATTAAGAAATTTAGTGTATTCAGAGGACATACAGAAAGAG 2480  
481 GAAATTTGTAGGAAAGCAATTAAGAAATTTAGTGTATTCAGAGGACATACAGAAAGAG 540  
2481 GGTGAAATGTGAGGCTTTGAGAAATTTGTCTATTTGACATATGAGGGTCAAGTACT 2540  
541 GGTGAAATGTGAGGCTTTGAGAAATTTGTCTATTTGACATATGAGGGTCAAGTACT 600  
2541 GCCTTACTTGGCCACAGTGGAAAGAGGACAGTACATTTGATGAATTTCTTGTGGACTC 2600  
601 GCCTTACTTGGCCACAGTGGAAAGAGGACAGTACATTTGATGAATTTCTTGTGGACTC 660  
2601 TGCCCACTTCTGATGGTTTGCATCTATATATGACACAGAGTCTCAGAAATAGATGAA 2660  
661 TGCCCACTTCTGATGGTTTGCATCTATATATGACACAGAGTCTCAGAAATAGATGAA 720  
2661 ATGTTTGAAGCAAGAAATGATTTGGCATTTGTCACAGTTAGATATACACTTTGATGTT 2720  
721 ATGTTTGAAGCAAGAAATGATTTGGCATTTGTCACAGTTAGATATACACTTTGATGTT 780  
2721 TTGACAGTACAGAAATTTATCAATTTTGGCTTCAATCAAGGATACCCAGCAACAAT 2780  
781 TTGACAGTACAGAAATTTATCAATTTTGGCTTCAATCAAGGATACCCAGCAACAAT 840  
2781 ATATAACAAGAGTGCAGAAAGTTTACTAGATTTAGACATGCAGACTATCAAGAATAAC 2840  
841 ATATAACAAGAGTGCAGAAAGTTTACTAGATTTAGACATGCAGACTATCAAGAATAAC 900  
2841 CAAGCTAAAAAATTAAGTGGTGCAGAAAGTTTACTAGATTTAGACATGCAGACTATGCCTTT 2900  
901 CAAGCTAAAAAATTAAGTGGTGCAGAAAGTTTACTAGATTTAGACATGCAGACTATGCCTTT 960  
2901 GGGAAACCAAGACTCTGCTAGATGAACCAACAGCTGGAATGACCCCTGCTTCGA 2960  
961 GGGAAACCAAGACTCTGCTAGATGAACCAACAGCTGGAATGACCCCTGCTTCGA 1020  
2961 CATATTGTATGGAATCTTTTAAATAACAGAAAGCCCAATCGGCTGACAGTGTTCAGTACT 3020  
1021 CATATTGTATGGAATCTTTTAAATAACAGAAAGCCCAATCGGCTGACAGTGTTCAGTACT 1080  
3021 CATTTATGATGAGTCAATCTTTGAGATGAGGAAAGCTGTGATATCAAGAAGATG 3080  
1081 CATTTATGATGAGTCAATCTTTGAGATGAGGAAAGCTGTGATATCAAGAAGATG 1140  
3081 CTGAAATGCTGTGCTTCAATGTTCTCAAGTAAATCGGGGATCGGCTACCGCTG 3140  
1141 CTGAAATGCTGTGCTTCAATGTTCTCAAGTAAATCGGGGATCGGCTACCGCTG 1200  
3141 AGCATGTACATAGACAAATATTTGTGCCACAGAAATCTCTTTCTTCACTGGTTAAACAACAT 3200

Db 1201 AGCATGTATAGACAAATATTTGTGCCACAGAAATCTCTTTCTTCACTGGTTAAACAACAT 1260  
Qy 3201 ATACCTGAGCTACTTTTATTAACAACAGAAATGACCAACAACTTGTATAGCTTGGCTTTC 3260  
Db 1261 ATACCTGAGCTACTTTTATTAACAACAGAAATGACCAACAACTTGTATAGCTTGGCTTTC 1320  
Qy 3261 AAGGACATGGACAAATTTTTCAG-----GTTTG 3287  
1321 AAGGACATGGACAAATTTTTCAGAAATGCTTGTATAGACAAAGGATGTTATTTGATGTTTG 1380  
3288 TTTTCTGCCCTAGACAGTCAATTTTAAATTTGGTGTCTTCTTATGGTGTTCATGAGC 3347  
1381 TTTTCTGCCCTAGACAGTCAATTTTAAATTTGGTGTCTTCTTATGGTGTTCATGAGC 1440  
3348 ACTTTTGAAGAGTATTTTAAAGCTAGAAAGTTGAAGCAGAAATGACCAAGCAGATAT 3407  
1441 ACTTTTGAAGAGTATTTTAAAGCTAGAAAGTTGAAGCAGAAATGACCAAGCAGATAT 1500  
3408 AGTGTATTTTACTCAGCAGCCACTGGAGGAAGAAATGGAAATCTTTTGTATGAAATG 3467  
1501 AGTGTATTTTACTCAGCAGCCACTGGAGGAAGAAATGGAAATCTTTTGTATGAAATG 1560  
3468 GAAACAGAGTACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTT 3527  
1561 GAAACAGAGTACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTT 1620  
3528 TGGAAACACAGATGATATACAAATGCTTCTTCTTACCTTGAACCTGCAACCTGCAAGT 3587  
1621 TGGAAACACAGATGATATACAAATGCTTCTTCTTACCTTGAACCTGCAACCTGCAAGT 1680  
3588 AAATCAGTGAGATCAGTGTCTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATG 3647  
1681 AAATCAGTGAGATCAGTGTCTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATG 1740  
3648 TTTTGGTTCATCACTCTTTTAAATGCTGTGGTTCCTCATCAAACTTGTTCACACTTA 3707  
1741 TTTTGGTTCATCACTCTTTTAAATGCTGTGGTTCCTCATCAAACTTGTTCACACTTA 1800  
3708 TATTTTCTTAAACCTGGAGACAAACACATAAATACAAACAAAGTCTGCTTCTTCAAAAT 3767  
1801 TATTTTCTTAAACCTGGAGACAAACACATAAATACAAACAAAGTCTGCTTCTTCAAAAT 1860  
3768 TCTGCTG 3774  
1861 TCTGCTG 1867  
RESULT 11  
AAS63176  
ID AAS63176 standard; cDNA; 1818 BP.  
XX AC AAS63176;  
XX DT 29-JAN-2002 (first entry)  
XX Human purified secretory polynucleotide #32.  
DE Human; purified secretory polypeptide; cell proliferative disorder; ss;  
XX cancer; immune system disorder; neurological disorder; mental disorder;  
KW motor neuron disorder; demyelinating disorder; neuromuscular disorder;  
KW central nervous system disorder; enzyme linked immunosorbent assay;  
KW ELISA; gene therapy.  
XX OS Homo sapiens.  
XX PN WO200162918-A2.  
XX PD 30-AUG-2001.  
XX PF 01-FEB-2001; 2001WO-US03465.  
XX PR 24-FEB-2000; 2000US-185215P.  
PR 24-FEB-2000; 2000US-185216P.

```
PR 16-MAY-2000; 2000US-205232P.
PR 17-MAY-2000; 2000US-205286P.
PR 17-MAY-2000; 2000US-205287P.
PR 17-MAY-2000; 2000US-205323P.
PR 17-MAY-2000; 2000US-205324P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;
PI Chen A, D'ea SA, Amshew S, Dahl CR, Dam TC, Daniele SE;
PI Dufour GE, Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL;
PI Liu TF, Roseberry AM, Rosen BH, Russo FD, Stockreher TK, Daffo A;
PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;
PI Cohen HJ, Hodgson DM, Lincoln SE;
XX
XX WPI; 2001-648217/74.
XX
XX Nucleic acids encoding secretory polypeptides, useful in genetic
XX diagnosis and therapy -
XX
XX Claim 1; Page 153-154; 237pp; English.
XX
XX Sequences AAS63145-AAS63223 represent DNA encoding purified secretory
XX polypeptides of the invention. The polypeptides and polynucleotides can
XX be used in the treatment, prevention and diagnosis of diseases associated
XX with inappropriate secretory protein expression. These diseases include
XX cell proliferative disorders such as atherosclerosis and psoriasis,
XX cancers such as leukaemia and melanoma, immune system disorders such as
XX asthma and diabetes mellitus, neurological disorders such as epilepsy and
XX Parkinson's disease, mental disorders such as schizophrenia and seasonal
XX affective disorder (SAD), motor neuron disorders such as amyotrophic
XX lateral sclerosis, demyelinating disorders such as multiple sclerosis,
XX central nervous system disorders such as mental retardation and
XX neurofibromatosis and neuromuscular disorders such as cerebral palsy and
XX muscular dystrophy. Target polynucleotides in a sample can be detected by
XX hybridising the sample with a probe sequence complementary to the target
XX polynucleotide, under conditions in which a hybridisation complex is
XX formed, and detecting the presence or absence of the complex. The
XX polypeptides may also be used as antigens in the production of antibodies
XX against secretory proteins and in assays to identify modulators of
XX protein expression and activity. The antibodies may also be used as
XX diagnostic agents for detecting the presence of the sequences of the
XX invention in samples e.g. by enzyme linked immunosorbent assay (ELISA).
XX
XX Sequence 1818 BP; 482 A; 347 C; 362 G; 627 T; 0 other;
XX
XX
XX Query Match 26.4%; Score 1722.8; DB 22; Length 1818;
XX Best Local Similarity 99.8%; Pred. No. 0;
XX Matches 1735; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
XX
XX 1 AAAATGTTGATATTTCTTAGCAGGCTCTCAACAGGTTAGGTTGAGTTCAGTTCATAGTTT 60
XX |||||
XX 81 AAAATGTTGATATTTCTTAGCAGGCTCTCAACAGGTTAGGTTGAGTTCAGTTCATAGTTT 140
XX |||||
XX 61 CTACCCACATCTTTGAACGTAGTGTCTCATTTTAGTTATTTTCAAAAACCTTTTGCAG 120
XX |||||
XX 141 CTACCCACATCTTTGAACGTAGTGTCTCATTTTAGTTATTTTCAAAAACCTTTTGCAG 200
XX |||||
XX 121 TACCTTTTGGTGTGCTTGTGTGTGCTTGCAGTGAACAGTCTGGATTTGGACAGTGT 180
XX |||||
XX 201 TACCTTTTGGTGTGCTTGTGTGTGCTTGCAGTGAACAGTCTGGATTTGGACAGTGT 260
XX |||||
XX 181 CTGTCTGTAGTTAGTTCAGTTTCTCAAGCCTTTGTCACTAATAGGTTGGATTTATGTATG 240
XX |||||
XX 261 CTGTCTGTAGTTAGTTCAGTTTCTCAAGCCTTTGTCACTAATAGGTTGGATTTATGTATG 320
XX |||||
XX 241 TCACAGCTTGGGAATATTATCAGGAATTAATAAACAACCTTTTAGAGTCTTTTCTTGAGCTC 300
XX |||||
XX 321 TCACAGCTTGGGAATATTATCAGGAATTAATAAACAACCTTTTAGAGTCTTTTCTTGAGCTC 380
XX |||||
XX 301 TCTTTCTATTGTTCCCTTCTTACTTTTGTGCTTCCCTGTGGCTGTGCTTCTATCCTCC 360
XX |||||
XX 381 TCTTTCTATTGTTCCCTTCTTACTTTTGTGCTTCCCTGTGGCTGTGCTTCTATCCTCC 440
XX |||||
QY 361 AGCCAGAGAGCTAGTGTATTTTCTCCATTTGTGTTTACACACTTGTGACGCTGCAACCCAC 420
DB 441 AGCCAGAGAGCTAGTGTATTTTCTCCATTTGTGTTTACACACTTGTGACGCTGCAACCCAC 500
QY 421 CATATCCAGGCCCCAATGGTAGAGGTAGAGAGAAAGCAAAAGGATTTGGCTCATCC 480
DB 501 CATATCCAGGCCCCAATGGTAGAGGTAGAGAGAAAGCAAAAGGATTTGGCTCATCC 560
QY 481 TCTTCAACAGTAGTTTCCATTTGAATAGAGAGAAAGGTTTTTCCCTGCTCAGAGTGTGGCT 540
DB 561 TCTTCAACAGTAGTTTCCATTTGAATAGAGAGAAAGGTTTTTCCCTGCTCAGAGTGTGGCT 620
QY 541 GCACTAGGCTTTTGTGTTTACTGTAGTCTGCGCCCTCTTACCATTGGATTTGCTATGTGGGG 600
DB 621 GCACTAGGCTTTTGTGTTTACTGTAGTCTGCGCCCTCTTACCATTGGATTTGCTATGTGGGG 680
QY 601 ATACAGGAGATTTCAAGAAAGAAAGATTTGCTATTTTCTACATTTCCCTGAGCATTT 660
DB 681 ATACAGGAGATTTCAAGAAAGAAAGATTTGCTATTTTCTACATTTCCCTGAGCATTT 740
QY 661 AAGACCTCCCTGCCCCATTCCTCAATTTCAAAGCTTAAGGCTTCTCTGGAGCTGCCCTGT 720
DB 741 AAGACTTCCCTTGGCCCATTTCCCTCAATTTCAAAGCTTAAGGCTTCTCTGGAGCTGCCCTGT 800
QY 721 GGGCGGTTGGGAGATACCAAGAGGAAAGATACCACTGTTGATATGGTGGTATTTCAA 780
DB 801 GGGCGGTTGGGAGATACCAAGAGGAAAGATACCACTGTTGATATGGTGGTATTTCAA 860
QY 781 ATTCTGGTCTACCTATTTTCAACATGCTTGTGTTTACTTTTTCAGAGCTGACAGATTTGCT 840
DB 861 ATTCTGGTCTACCTATTTTCAACATGCTTGTGTTTACTTTTTCAGAGCTGACAGATTTGCT 920
QY 841 CCATGCAATTTCTGCTCAGTTTCTTAAAGAGAGACAGCTTGGAGTATGCTTAATCCATCTTAC 900
DB 921 CCATGCAATTTCTGCTCAGTTTCTTAAAGAGAGACAGCTTGGAGTATGCTTAATCCATCTTAC 980
QY 901 CTGGGACTCAAAACAGCTGCTTATTTTGGCGTTAAATAATACATGCAAGTTTACTGCGTGGC 960
DB 981 CTGGGACTCAAAACAGCTGCTTATTTTGGCGTTAAATAATACATGCAAGTTTACTGCGTGGC 1040
QY 961 TCCGGGTTTGTGTTGTTGTTTTCCT - CTTTAAATAGTTTATTTTCAAGAAACATGTCACCT 1019
DB 1041 TCCGGGTTTGTGTTGTTGTTTTCCTTTTAAATAGTTTATTTTCAAGAAACATGTCACCT 1100
QY 1020 GCNATTTAGGAGGTAGAGTTTGGAGACAGACACACTTCTACTGAGAAATTAATTA 1079
DB 1101 GCAATTTAGGAGGTAGAGTTTGGAGACAGACACACTTCTACTGAGAAATTAATTA 1160
QY 1080 ATTTAAATGAGAACCAAAAGAGTAGTGTTCAGGAAATTTCTTTTCCACTATTTTATTA 1139
DB 1161 ATTTAAATGAGAACCAAAAGAGTAGTGTTCAGGAAATTTCTTTTCCACTATTTTATTA 1220
QY 1140 TTTTGGTTTAAATTAATTTAGCATGTCATCCAAATTAAGAAATATGAAGAGTGCCTAAT 1199
DB 1221 TTTTGGTTTAAATTAATTTAGCATGTCATCCAAATTAAGAAATATGAAGAGTGCCTAAT 1280
QY 1200 ATGAAACTCAATCTATGACAAAGTTTACTCTTTTAAATCTAATTTCTTGGATATCTCCA 1259
DB 1281 ATGAAACTCAATCTATGACAAAGTTTACTCTTTTAAATCTAATTTCTTGGATATCTCCA 1340
QY 1260 GTGACTTAATTTACAGAGCAGCATGTCAGAAAGTGTCTACTCATCTACCTGATGTC 1319
DB 1341 GTGACTTAATTTACAGAGCAGCATCATGAGAAAGTGTCTACTGATCATCTACCTGATGTC 1400
QY 1320 ATAAATTTAGGAAATATACAAATGAAAGAAATGTTAAACATCCAGTCTCTCTAAGCCG 1379
DB 1401 ATAAATTTAGGAAATATACAAATGAAAGAAATGTTAAACATCCAGTCTCTCTAAGCCG 1460
QY 1380 AGCAACTTTTGTAGGTGTGTTTCAAAGACTCCATGTCCTATGAACTTCGTTTTTTTCT 1439
DB 1461 AGCAACTTTTGTAGGTGTGTTTCAAAGACTCCATGTCCTATGAACTTCGTTTTTTTCT 1520
```

QY 1440 GATATGATCCAGTATCTTCTATTTATATGGAATCAAGAGCTGGCTGTTCAAAATCATGT 1499  
DB |||||  
QY 1521 GATATGATCCAGTATCTTCTATTTATATGGAATCAAGAGCTGGCTGTTCAAAATCATGT 1580  
DB |||||  
QY 1500 GAGCTGCTCAGTACTGGTCTCAGGTTTACAGTTTACAGATCCATAGATGCTGCC 1559  
DB |||||  
QY 1581 GAGCTGCTCAGTACTGGTCTCAGGTTTACAGTTTACAGATCCATAGATGCTGCC 1640  
DB |||||  
QY 1560 ATTATACAGTTGAAGACCAATGTTCTCTTTGGAAGAGCTGGCTCACTAAAGCTGTT 1619  
DB |||||  
QY 1641 ATTATACAGTTGAAGACCAATGTTCTCTTTGGAAGAGCTGGCTCACTAAAGCTGTT 1700  
DB |||||  
QY 1620 ATTATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGAGTAATTTTAAATA 1679  
DB |||||  
QY 1701 ATTATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGAGTAATTTTAAATA 1760  
DB |||||  
QY 1680 TACTAGTTATAGATTTTACCTTTTGGATACCTTTTGGCAATTCATATCGTAGCAG 1737  
DB |||||  
QY 1761 TACTAGTTATAGATTTTACCTTTTGGATACCTTTTGGCAATTCATATCGTAGCAG 1818  
DB |||||

RESULT 12  
ABK35706  
ID ABK35706 standard; cDNA; 1548 BP.

XX ABK35706;

DT 08-MAY-2002 (first entry)

XX cDNA sequence #97 encoding novel human secreted protein.

XX Human secreted protein; hyperproliferative disorder; autoimmune disorder;  
XX immune deficiency disorder; blood disorder; inflammatory disorder;  
XX infectious disorder; allergic condition; neurodegenerative disorder;  
XX liver fibrosis; coagulation disorder; gene therapy; antimicrobial;  
XX tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.

OS Homo sapiens.

XX W0200177289-A2.

XX 18-OCT-2001.

XX 29-MAR-2001; 2001WO-US10232.

XX 06-APR-2000; 2000US-195605P.

XX (GEMY ) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
XX Marberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GG;  
XX Clark HF, Fechtel K, Howes SH, Resnick RJ, Gulukota K, Graham JR;  
XX WPI; 2002-179322/23.

XX Six hundred and twenty three polynucleotides derived from a variety of  
XX human tissue sources which encode secreted proteins, useful for  
XX treating immune deficiencies and disorders such as autoimmune disorders

XX Claim 1; Page 126; 393pp; English.

XX The present invention relates to the isolation of novel cDNA sequences  
XX which encode human secreted proteins. The cDNA sequences have been  
XX derived from a variety of human tissues. The invention also provides  
XX a method for producing proteins from these polynucleotide sequences.  
XX The proteins are useful for identifying compounds that modulate their  
XX activity and production. The sequences of the invention are  
XX useful for treating diseases such as hyperproliferative disorders  
XX (e.g. cancer), immune deficiency disorders (e.g. severe combined  
XX immunodeficiency (SCID)), autoimmune disorders (e.g. multiple  
XX sclerosis), blood disorders (e.g. thrombocytopaenia), inflammatory  
XX disorders (e.g. arthritis), infectious disorders (e.g. hepatitis),

CC allergic conditions (e.g. asthma), neurodegenerative disorders (e.g.  
CC Alzheimer's disease), liver fibrosis, coagulation disorders (e.g.  
CC haemophilia), and tumours. The polynucleotide sequences of the  
CC invention are also useful in gene therapy. ABK35610-ABK36232 represent  
CC the cDNA sequences of the invention that encode for novel human  
CC secreted proteins.

XX Sequence 1548 BP; 466 A; 273 C; 268 G; 541 T; 0 other;

Query Match 23.7%; Score 1543.8; DB 24; Length 1548;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1545; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3386 AGAAATTCACCAAGCAGATTTAGTGTATTTACTCAGCAGCCACTCGAGGAAGAATGGA 3445  
DB 1 AGAAATTCACCAAGCAGATTTAGTGTATTTACTCAGCAGCCACTCGAGGAAGAATGGA 60

QY 3446 TTCAAAATCTTTTGATGAAATGGAACAGAGCTTACTTCTTCTGAAACCAAGGCTTC 3505  
DB 61 TTCAAAATCTTTTGATGAAATGGAACAGAGCTTACTTCTTCTGAAACCAAGGCTTC 120

QY 3506 TCTAGTGAGCACCATGAGCCTTTGGAACACAGAGTATACAATAGCAAAAGTTCATTT 3565  
DB 121 TCTAGTGAGCACCATGAGCCTTTGGAACACAGAGTATACAATAGCAAAAGTTCATTT 180

QY 3566 CTTTACCTTGAACCGTGAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTT 3625  
DB 181 CTTTACCTTGAACCGTGAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTT 240

QY 3626 TTTCACAGTTCAGATTTTATGTTTGTTCATCCTCTTTTAAAAATGCTGTGGTTC 3685  
DB 241 TTTCACAGTTCAGATTTTATGTTTGTTCATCCTCTTTTAAAAATGCTGTGGTTC 300

QY 3686 CATCAAACTGTTTCAGACTTATATTTCTAAACCTGGAGACCAACACATAATACAA 3745  
DB 301 CATCAAACTGTTTCAGACTTATATTTCTAAACCTGGAGACCAACACATAATACAA 360

QY 3746 AACCAAGTCTGCTTCTTCAAAATCTGCTGACTCAGATATCAGTGATCTTATTAGCTTTT 3805  
DB 361 AACCAAGTCTGCTTCTTCAAAATCTGCTGACTCAGATATCAGTGATCTTATTAGCTTTT 420

QY 3806 CACAAGCCAGAACATAATGGTGACGATTAATGACAGTGATATGATATCGTGGCTCC 3865  
DB 421 CACAAGCCAGAACATAATGGTGACGATTAATGACAGTGATATGATATCGTGGCTCC 480

QY 3866 CCATAGTGGGCTTTAAATGTGATGATCTCAGAAAGGACTATGTTTTCAGAGCTGTTT 3925  
DB 481 CCATAGTGGGCTTTAAATGTGATGATCTCAGAAAGGACTATGTTTTCAGAGCTGTTT 540

QY 3926 CAACAGTACTATGGTTTATTTCTTTACCTATATTAGTGAATATCATTAGTAATCTATCT 3985  
DB 541 CAACAGTACTATGGTTTATTTCTTTACCTATATTAGTGAATATCATTAGTAATCTATCT 600

QY 3986 TTATCATTTAAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATCTTTTCAAGAAAT 4045  
DB 601 TTATCATTTAAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATCTTTTCAAGAAAT 660

QY 4046 TACTGATATAGTTTAAATGTGATGATCTCAGAAAGGACTATGTTTTCAGAGCTGTTT 4105  
DB 661 TACTGATATAGTTTAAATGTGATGATCTCAGAAAGGACTATGTTTTCAGAGCTGTTT 720

QY 4106 TACTGCAATGCCACCTTACTTTTCCATGGAATATCAGAGATCATAGATCAAAAGCTTA 4165  
DB 721 TACTGCAATGCCACCTTACTTTTCCATGGAATATCAGAGATCATAGATCAAAAGCTTA 780

QY 4166 TACTCAACTTAACTTTTTCAGGCTTTTCCATCTCCATATTTGGATTTGGACAGCTGTTT 4225  
DB 781 TACTCAACTTAACTTTTTCAGGCTTTTCCATCTCCATATTTGGATTTGGACAGCTGTTT 840

QY 4226 TGATATCCCTTATTTTATTCATTTTATTTGATGCTAGGAAGCTTACTGGCATTTC 4285  
DB 841 TGATATCCCTTATTTTATTCATTTTATTTGATGCTAGGAAGCTTACTGGCATTTC 900

Qy	4286	TTATGATTATATTTTATATCTGTAAAGTTCCTTCTGCTGGTCTTTTGGCTTATTTGTTA	4345
Db	901	TTATGATTATATTTTATATCTGTAAAGTTCCTTCTGCTGGTCTTTTGGCTTATTTGTTA	960
Qy	4346	TGTTCCATCAGTTATCTGTCTTACTTATATGCTTCTTTTACCTTTTAAAGAAATTTTAA	4405
Db	961	TGTTCCATCAGTTATCTGTCTTACTTATATGCTTCTTTTACCTTTTAAAGAAATTTTAA	1020
Qy	4406	TACCAAGAAATTTTGGTCAATTTATCTTCTGTGGCAGCGTTCGCTTGTATTCGAATCAC	4465
Db	1021	TACCAAGAAATTTTGGTCAATTTATCTTCTGTGGCAGCGTTCGCTTGTATTCGAATCAC	1080
Qy	4466	TGAATAAATCTTCTTTATGGATACACAAATTCGAATTCCTTCATATGCCCTTTTGTAT	4525
Db	1081	TGAATAAATCTTCTTTATGGATACACAAATTCGAATTCCTTCATATGCCCTTTTGTAT	1140
Qy	4526	CATCATTCATCTATCCACTTCTAGTTCGCTGATTTCTTCATTAAGATTTTCTTGGAA	4585
Db	1141	CATCATTCATCTATCCACTTCTAGTTCGCTGATTTCTTCATTAAGATTTTCTTGGAA	1200
Qy	4586	GAATGTACGAAAAAATGTGACACCTATATATCCATGGATAGGCTTTTCAGTAGCTGTAT	4645
Db	1201	GAATGTACGAAAAAATGTGACACCTATATATCCATGGATAGGCTTTTCAGTAGCTGTAT	1260
Qy	4646	ATCGCTTACCTGAGTGTGTACTGTGATTTTCTTCAATATCTATGAGAAAAAATA	4705
Db	1261	ATCGCTTACCTGAGTGTGTACTGTGATTTTCTTCAATATCTATGAGAAAAAATA	1320
Qy	4706	TGAGGCGAGATCAATAGAAAGATCCCTTTTTCAGAAACCTTTTCAAGAGTCTTAAAA	4765
Db	1321	TGAGGCGAGATCAATAGAAAGATCCCTTTTTCAGAAACCTTTTCAAGAGTCTTAAAA	1380
Qy	4766	TAGGAAGCTTCCAGAACCCAGACCAATGAGATGGAAGTCAAGTCTCAAGCTGAAG	4825
Db	1381	TAGGAAGCTTCCAGAACCCAGACCAATGAGATGGAAGTCAAGTCTCAAGCTGAAG	1440
Qy	4826	ACTAAAGGTCAAAGAGCTGATGGTGGCCAGTGTGTGGAGGAAACCAATCCATTTAGT	4885
Db	1441	ACTAAAGGTCAAAGAGCTGATGGTGGCCAGTGTGTGGAGGAAACCAATCCATTTAGT	1500
Qy	4886	CAGCAATTTCCATTAAGAAATATGATGACAAAGATTTTCTTCTTTC 4933	
Db	1501	CAGCAATTTCCATTAAGAAATATGATGACAAAGATTTTCTTCTTTC 1548	
RESULT 13			
ID	AAC81717		
AC	AAC81717	standard; DNA; 1506 BP.	
AC	AAC81717;		
DT	26-FEB-2001	(first entry)	
DE	Human secreted protein coding sequence SEQ ID NO: 18.		
XX	Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;		
KW	allergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;		
KW	vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic;		
KW	cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;		
KW	neurological disease; infection; human; secreted protein; ss.		
OS	Homo sapiens.		
XX	W0200061627-A1.		
XX	19-OCT-2000.		
PF	06-APR-2000; 2000WO-US09067.		
XX	09-APR-1999; 99US-0128697.		
PR	20-JAN-2000; 2000US-0176929.		
XX	(HUMA-) HUMAN GENOME SCI INC.		

PA	(ROSE/) ROSEN C A.		
XX	Rosen CA, Ruben SM, Komatsoulis G;		
XX	WPI; 2000-647419/62.		
DR	P-PSDB; AAB45393.		
PT	New nucleic acid molecules encoding 49 human secreted proteins for		
PT	diagnosing, preventing, treating or ameliorating medical conditions and		
PT	used as food additives or preservatives -		
XX	Claim 1; Page 429; 521pp; English.		
PS	The invention relates to the isolation of genes AAC81710-C81758 encoding		
XX	49 human secreted proteins AAB45386-B54534. The genes can be used to		
CC	generate fusion proteins by linking to the gene for the human		
CC	immunoglobulin G Fc portion (SEQID1) for increasing the stability of		
CC	the fusion protein as compared to the human protein only. The genes and		
CC	proteins are useful for preventing, ameliorating or treating medical		
CC	conditions, e.g. by protein or gene therapy. The genes are isolated		
CC	from a range of human tissues disclosed in the specification. The		
CC	nucleic acids, proteins, antibodies and (ant)agonists are useful in		
CC	the diagnosis, treatment and prevention of: (a) cancer, e.g. breast		
CC	and ovarian cancer, and other cancers of the adrenal gland, bone, bone		
CC	marrow, breast, gastrointestinal tract, liver, lung, or urogenital;		
CC	(b) immune disorders e.g. Addison's disease, allergies, autoimmune		
CC	haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's		
CC	disease, multiple sclerosis, rheumatoid arthritis and ulcerative		
CC	colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)		
CC	wound healing; (e) neurological diseases e.g. cerebral anoxia and		
CC	epilepsy; and (f) infectious diseases such as viral, bacterial, fungal		
CC	and parasitic infections.		
XX	Sequence 1506 BP; 509 A; 218 C; 290 G; 485 T; 4 other;		
SQ	Query Match 20.2%; Score 1319; DB 21; Length 1506;		
	Best Local Similarity 92.1%; Pred. No. 8.7e-272;		
	Matches 1460; Conservative 1; Mismatches 4; Indels 121; Gaps 2;		
Qy	4941	GTAAAGAAAGTGGCAACTAAATACATCTCTTCTGTGTGTAAGAAAGAGAGATCTTAGGA	5000
Db	1	GTAAAGAAAGTGGCAACTAAATACATCTCTTCTGTGTGTAAGAAAGAGAGATCTTAGGA	60
Qy	5001	CTATTGGGTCCAAATGGTGTGCAAAAGCAATTAATTAATTTCTGGTGTGTATTT	5060
Db	61	CTATTGGGTCCAAATGGTGTGCAAAAGCAATTAATTAATTTCTGGTGTGTATTT	120
Qy	5061	GAAACCACTTCAGGCCAGGTATTTTGGAGATTTATTTCTTCAGAGACAAGTGAAGATGAT	5120
Db	121	GAAACCACTTCAGGCCAGGTATTTTGGAGATTTATTTCTTCAGAGACAAGTGAAGATGAT	180
Qy	5121	GATTCACCTGAAGTGTATGGTTACTGTCTCAGATAAACCTTTTGGCCAGATCTACA	5180
Db	181	GATTCACCTGAAGTGTATGGTTACTGTCTCAGATAAACCTTTTGGCCAGATCTACA	240
Qy	5181	TTGCAGAAACATTTTGAATTTTATGAGCTGTCAAAAGGAATGAGTCAAGTGCATCAAA	5240
Db	241	TTGCAGAAACATTTTGAATTTTATGAGCTGTCAAAAGGAATGAGTCAAGTGCATCAAA	300
Qy	5241	GAAAGTCATAGTGAATTAACACATGACCTTGAATTTTAAAGAAACATCTTCAGAAAGCTGA	5300
Db	301	GAAAGTCATAGTGAATTAACACATGACCTTGAATTTTAAAGAAACATCTTCAGAAAGCTGA	360
Qy	5301	AGAATACTACCTGCAGGAATCAACGAAAGTTGTGTTTGTCTTCTAGTATGCTAGGAAT	5360
Db	361	AGAATACTACCTGCAGGAATCAACGAAAGTTGTGTTTGTCTTCTAGTATGCTAGGAAT	420
Qy	5361	CCTCAGATTACTTTTGTAGATGAACCATCTACAGGTATGATCCCAAGGCAAGCAAGCAC	5420
Db	421	CCTCAGATTACTTTTGTAGATGAACCATCTACAGGTATGATCCCAAGGCAAGCAAGCAC	480
Qy	5421	ATGTGGCGAGCAATTCGAACTGCATTTTAAAAACAGAAAGCGGCTGCTATTCTTGACCACT	5480

Db 481 ATGT----- 484  
QY 5481 CACTATATGAGGAGGAGGAGGCTGTCTGTGATCGAGTAGCTATCATGTGTCTGGCAG 5540  
Db 485 ----- 484  
QY 5541 TTAAGATGTATCGAACAGTACACATCTAAAGAGTAAATTTGGAAAAAGGCTACTTTTGG 5600  
Db 485 -----GATGTATCGAACAGTACACATCTAAAGAGTAAATTTGGAAAAAGGCTACTTTTGG 540  
QY 5601 GAAATTTAAATTTGAAGGACTCGATAGAAAACCTAGAAAGTAGACCGCTTCAAAGAGAAAT 5660  
Db 541 GAAATTTAAATTTGAAGGACTCGATAGAAAACCTAGAAAGTAGACCGCTTCAAAGAGAAAT 600  
QY 5661 CAGTATATTTCCCAATTCGAAGCGTCAGAAAGTTTTCTTCTATTTTGGCTTATAA 5720  
Db 601 CAGTATATTTTCCCAATTCGAAGCGTCAGAAAGTTTTCTTCTATTTTGGCTTATAA 660  
QY 5721 ATTCTTAAGGAAGATGTTCAGTCCCTTTTCAATCTTTTTTAAAGCTGGAAGAGCTAAA 5780  
Db 661 ATTCTTAAGGAAGATGTTCAGTCCCTTTTCAATCTTTTTTAAAGCTGGAAGAGCTAAA 720  
QY 5781 CATGCTTTTCCCAATTCGAAGCTATAGCTTTTCTCAAGCAACATGGAACAGGTTTTTGT 5840  
Db 721 CATGCTTTTCCCAATTCGAAGCTATAGCTTTTCTCAAGCAACATGGAACAGGTTTTTGT 780  
QY 5841 GAACTCACTAAGAACCAAGAGGAGGAGATAGTTGTGGAACCTTTAAACAGCACACT 5900  
Db 781 GAACTCACTAAGAACCAAGAGGAGGAGATAGTTGTGGAACCTTTAAACAGCACACT 840  
QY 5901 TGGTGGGAACGAAACACAGAGAGATAGTAGTATTTTGAATTTGTATTGTTGGCTGTCT 5960  
Db 841 TGGTGGGAACGAAACACAGAGAGATAGTAGTATTTTGAATTTGTATTGTTGGCTGTCT 900  
QY 5961 TACTGGGACTCTTTCTTTTCACTTAATTTAACTTTGTTGTTTAAAGTTTTTATG 6020  
Db 901 TACTGGGACTCTTTCTTTTCACTTAATTTAACTTTGTTGTTTAAAGTTTTTATG 960  
QY 6021 AATGTTAATCTGGAGAACCAAGAGGAGGAGATAGTTGTGGAACCTTTAAACAGCACACT 6080  
Db 961 AATGTTAATCTGGAGAACCAAGAGGAGGAGATAGTTGTGGAACCTTTAAACAGCACACT 1020  
QY 6081 CTGTGGTGTGTGTTTCTTTTCACTTAATTTAACTTTGTTGTTTAAAGTTTTTATG 6140  
Db 1021 CTGTGGTGTGTGTTTCTTTTCACTTAATTTAACTTTGTTGTTTAAAGTTTTTATG 1080  
QY 6141 ATGTTTGTATTGAAGTATATGAACTATATAGTTTGTATGTCATCTTTTCCACTTCAG 6200  
Db 1081 ATGTTTGTATTGAAGTATATGAACTATATAGTTTGTATGTCATCTTTTCCACTTCAG 1140  
QY 6201 AAACAGTGTCTTGAATTTGATTTAAAGAAATGTAATAGATAGTTTATTTTAAAG 6260  
Db 1141 AAACAGTGTCTTGAATTTGATTTAAAGAAATGTAATAGATAGTTTATTTTAAAG 1200  
QY 6261 TTATCTTTTAAAGTTTATGCTATCTTTTAAAGTAACTAGTATGTTTCCAACTTAAATAA 6320  
Db 1201 TTATCTTTTAAAGTTTATGCTATCTTTTAAAGTAACTAGTATGTTTCCAACTTAAATAA 1260  
QY 6321 AACTAATACATACTAATATGATAGAAAAGATACATAAAGCAATGTGAAAGTTTCTTGT 6380  
Db 1261 AACTAATACATACTAATATGATAGAAAAGATACATAAAGCAATGTGAAAGTTTCTTGT 1320  
QY 6381 CTCCTTTTAAATTTCTAAAAAGGCACTTTGATGGAAGTTGTCATCCGTTAAAGCTGAA 6440  
Db 1321 CTCCTTTTAAATTTCTAAAAAGGCACTTTGATGGAAGTTGTCATCCGTTAAAGCTGAA 1380  
QY 6441 GTGTAAGCACTTAGGAATCTCAATATATAGATTTTGAAGAAAGTTATATCCACT-AGGTGG 6499  
Db 1381 GTGTAAGCACTTAGGAATCTCAATATATAGATTTTGAAGAAAGTTATATCCACTAAGTGG 1440  
QY 6500 CAGTCATTGTATCAATAAAGTGAAT 6525  
Db 1441 CAGTCATTGTATCAATAAAGTGAAT 1466

RESULT 14  
ABN96909  
ID ABN96909 standard; DNA; 1346 BP.  
XX  
AC ABN96909;  
XX  
DT 13-AUG-2002 (first entry)  
XX  
DE Gene #3407 used to diagnose liver cancer.  
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
KW metastatic liver tumour; cytostatic; expression profile; disease state;  
KW disease progression; drug toxicity; drug efficacy; drug metabolism.  
XX  
OS Homo sapiens.  
XX WO200229103-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 02-OCT-2001; 2001WO-US30589.  
XX  
PR 02-OCT-2000; 2000US-237054P.  
XX  
(GENE-) GENE LOGIC INC.  
PA Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
PI WPI; 2002-426119/45.  
XX  
DR Diagnosing and detecting the progression of liver cancer.  
XX hepatocellular carcinoma or metastatic liver tumor in a patient,  
PT involves detecting the level of expression of two or more genes in a  
PT liver tissue sample  
XX  
PS Claim 1; SEQ ID NO 3407; 298pp; English.  
XX  
CC The invention relates to a novel method for diagnosing and detecting the  
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
CC tumour in a patient, and differentiating metastatic liver cancer from  
CC hepatocellular carcinoma in a patient, involving detecting the level of  
CC expression of two or more genes represented in ABN93503-ABN97455 in a  
CC tissue sample. The method of the invention has hepatotropic, and  
CC cytostatic activity. The method is useful for diagnosing and detecting  
CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
CC liver carcinoma in a patient. The method is useful for identifying  
CC expression profiles which serve as useful diagnostic markers as well as  
CC markers that can be used to monitor disease states, disease progression,  
CC drug toxicity, drug efficacy and drug metabolism.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1346 BP; 451 A; 206 C; 290 G; 395 T; 4 other;  
Query Match 18.1%; Score 1182; DB 24; Length 1346;  
Best Local Similarity 96.5%; Pred. No. 1.5e-242;  
Matches 1306; Conservative 0; Mismatches 29; Indels 19; Gaps 9;  
QY 4786 CAGCAATCAGATGAGATGAAGATGTCAAAGCTGAAAGACTTAAAGGTCAAAGAGCTGA 4845  
Db 1 CAGACATCAGATGAGATGAGATGTCAAAGCTGAAAGACTTAAAGGTCAAAGAGCTGA 60  
QY 4846 TGGTTTGCAGTGTGTGTGAGGAAACCATCCATTATGTCAGCAATTTGCATAAAGAT 4905  
Db 61 TGGTTTGCAGTGTGTGTGAGGAAACCATCCATTATGTCAGCAATTTGCATAAAGAT 120  
QY 4906 ATGATCAGCAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAGTGGCAACTAAATACA 4965  
Db 121 ATGATCAGCAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAGTGGCAACTAAATACA 180

```
Qy 4966 TCTCTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGGTCCAAATGGTGTGGCA 5025
Db 181 TCTCTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGGTCCAAATGGTGTGGCA 240
Qy 5026 AAAGCACAATTAATTAATTTCTGGTGGTGATATTGAACCACTTCAGGCCAGGTATTTT 5085
Db 241 AAAGCACAATTAATTAATTTCTGGTGGTGATATTGAACCACTTCAGGCCAGGTATTTT 300
Qy 5086 TAGGAGATTAATTTCTCAGACACAAGTGAAGATGATGATTCACCTGAAGTGTATGGTTACT 5145
Db 301 TAGGAGATTAATTTCTCAGACACAAGTGAAGATGATGATTCACCTGAAGTGTATGGTTACT 360
Qy 5146 GTCTCAGATAAACCCCTTTGTGGCCAGATACTACATTCAGGAGAACATTTTGAATTTATG 5205
Db 361 GTCTCAGATAAACCCCTTTGTGGCCAGATACTACATTCAGGAGAACATTTTGAATTTATG 420
Qy 5206 GAGCTGTCAAGGAATGAGTGCAGATGACATGAAGAAGTCATAAGTCGAATTAACACATG 5265
Db 421 GAGCTGTCAAGGAATGAGTGCAGATGACATGAAGAAGTCATAAGTCGAATTAACACATG 480
Qy 5266 CACTTGATTTAAAGAACATCTTTCAGAAAGACTGTAAAGAAACTACCTGCAGGAATCAAC 5325
Db 481 CACTTGATTTAAAGAACATCTTTCAGAAAGACTGTAAAGAAACTACCTGCAGGAATCAAA 540
Qy 5326 GAAAGTTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTAATTTCTAGATGAAC 5385
Db 541 CGAAAGTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTAATTTCTAGATGAAC 600
Qy 5386 CATCTACAGTATGGATCCCAA--GCCAAACAGCAGATGTGGGAGCAATTCGAATGC 5443
Db 601 CATCTACAGTCTGGATCCCAAATGCCAAACATGCACATGTGGCATGCAATTCGAATGC 660
Qy 5444 ATTTAAAGAACAGAAAGCGGCTGCTATCTTGACCACTCAGTATGAGGAGCAGAGGC 5503
Db 661 AT-----NNAGCGGCTGCTATCTTGACCACTCAGTATGAGGAGCAGAGGC 711
Qy 5504 TGCTGTGATCGAGTAGTATCATGTGTCTGGGCACTTAAGATGTATCGGAAACAGTACA 5563
Db 712 TGCTGTGATCGAGTAGTATCATGTGTCTGGGCACTTAAGATGTATCGGAAACAGTACA 771
Qy 5564 ACATCTAAGAGTAATTTGGAAAAAGCTACTTTTGGAAATTAATTTGAA--GGACTGGA 5622
Db 772 ACATCTAAGAGTAATTTGGAAAAAG--NACTTTTGGAAATTAATTTGAAAGCACTGGA 829
Qy 5623 TAGAAACCTTAGAG--TAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCA 5681
Db 830 TAGAAACCTTAGAGCTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCA 889
Qy 5682 AGCGTTCAGAAAGTTTTTCTTCTATTTTGGCTTTATAAAAATTCCTAAGGAAGATGTTTCA 5741
Db 890 AGCGTCA--GAAAGTTTTTCTTCTATTTTGGCTTCTAANAATTAATTAAGGAAGATGTTTCA 948
Qy 5742 TCCCTTTCAAAATTTTTTAAAGTGGAAAGAGCTGAAACATGCTTTTGGCCATTTGAAGAA 5801
Db 949 TCCCTTTCCCAATCTTTTAAAGTGGAAAGAGCTGAAAGCAATGCTTTTGGCCATTTGAAGAA 1007
Qy 5802 TATAGCTTTTCTCAAGCAATTTGGAACAGTTTTTGTAGAACTCACTAAGAACAGAG 5861
Db 1008 TATAGC--TTTCTCAAGCAATTTGGAACAGTTTTTGTAGAACTCACTAAGAACAGAG 1066
Qy 5862 GAGGAATATAGTGTGAACTTTTAAACAGCACACTTTGGTGGGAAACGAAACAAGAA 5921
Db 1067 GAGGAATATAGTGTGAACTTTTAAACAGCACACTTTGGTGGGAAACGAAACAAGAA 1125
Qy 5922 GATAGTAGTATTTTGAATTTGATTTGTTGGTCTGCTTACTGGGACTCTTTCTTTTT 5981
Db 1126 GATAGTAGTATTTTGAATTTGATTTGTTGGTCTGCTTACTGGGACTCTTTCTTTTT 1185
Qy 5982 CACTTAATTTTAACTTTGTTTAAAGAGTTTTTATTTGGAATGTTAATCGAGAACCAAG 6041
Db 1186 CACTTAATTTTAACTTTGTTTAAAGAGTTTTTATTTGGAATGTTAATCGAGAACCAAG 1245
Qy 6042 AAGCAGCTTGAATTTTTTCTAAGCTCCTTAATTAATGAAATGCTGTGTTGTTGCTTT 6101
```

```
Db 1246 AAGCAGCTTGAATTTTTTCTAAGCTCCTTAATTAATGAAATGCTGTGTTGCTTT 1305
Qy 6102 TTCTTTAAATAAACCCTATCTATAATTAAGTGAA 6135
Db 1306 TTCTTTAAATAAACCCTATCTATAATTAAGTGAA 1339

RESULT 15
ABL63763
ID ABL63763 standard; DNA; 1346 BP.
XX
AC ABL63763;
XX
DT 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:2100.
XX
KW Human; cancer; colon; breast; kidney; ovary; oesophagus; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 03-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
```



